(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 May 2001 (31.05.2001)

PCT

(10) International Publication Number
WO 01/38489 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US00/32227

(22) International Filing Date:
22 November 2000 (22.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/448,215 23 November 1999 (23.11.1999) US

(71) Applicant: **CORNELL RESEARCH FOUNDATION, INC.** [US/US]; Suite 105, 20 Thornwood Drive, Ithaca, NY 14850 (US).

(72) Inventors: **YODER, Olen, C.**; 44 Hungerford Road, Ithaca, NY 14850 (US). **TURGEON, Barbara, C.**; 608 Mitchell Street, Ithaca, NY 14850 (US). **LU, Shun-Wen**; 604 Winston Court, Apt. #4, Ithaca, NY 14850 (US).

(74) Agents: **GOLDMAN, Michael, L.** et al.; Nixon Peabody LLP, Clinton Square, P.O. Box 31051, Rochester, NY 14603 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/38489 A2

(54) Title: PEPTIDE SYNTHETASE GENE CPS1

(57) Abstract: The present invention relates to genes cloned from the plant pathogens *Cochliobolus heterostrophus, Alternaria solani, Fusarium graminearium,* and *Pyrenophora teres*, that encode a CPS1 peptide synthetase required for fungal pathogenesis. The nucleic acid molecules in a vector, a host cell, or a plant is also disclosed. The invention further provides a protein or polypeptide encoded by the CPS1 genes. Other aspects of the invention relate to a method of imparting disease resistance to a plant by overexpressing a protein of the present invention in a plant and a method for identifying inhibitors of a CPS1 protein in a sample.

PEPTIDE SYNTHETASE GENE CPS1

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation in part application of U.S. Application Serial No. 09/448,215, filed November 23, 1999.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

The present invention was made with support from the United States Government under Grant No. 96-35303-3198 from the USDA/NRI. The United States Government may have certain rights in the invention.

FIELD OF THE INVENTION

The present invention relates to nucleic acid molecules encoding peptide synthetase homologs of *Cochliobolus heterostrophus*, *Pyrenophora teres*, *Fusarium graminearium*, and *Alternaria solani* and uses thereof.

BACKGROUND OF THE INVENTION

There are approximately 30 species included in the genus *Cochliobolus*, nearly all of which are pathogens of wild grasses or cereals (Yoder et al, "*Cochliobolus spp.* And Their Host-Specific Toxins, in Carroll, eds., The Mycota Vol. 5: Plant Relationships, Part A, Berlin:Springer-Verlag, pp. 145-166 (1997)). *Cochliobolus heterostrophus* represents the most widely distributed species in the genus and can be found in many tropical and subtropical areas in the world. As a natural pathogen of corn, *C. heterostrophus* causes a disease frequently called leaf spot of maize in the old literature (Drechsler, "Leafspot of Maize Caused by *Ophiobolus Heterostrophus* n. sp., The Ascigerous Stage of a *Helminthosporium* Exhibiting Bipolar Germination," J. Agr. Res., 31:701-726 (1925); Drechsler, "Phytopathological and Taxonomic Aspects of *Ophiobolus, Pyrenophora, Helminthosporium*, and a New Genus, *Cochliobolus*," Phytopathol., 24:953-983 (1934); Yu, "Studies on *Helminthosporium* Leaf Spot of Maize," 3:273-318 (1933); Orillo, "Leafspot of Maize Caused by *Helminthosporium maydis*," 36:327-395 (1952). In the United States, *C. heterostrophus* is usually found in the warmer southern states, thus, the disease is commonly known as Southern Corn Leaf Blight (Hooker, "Cytoplasmic Susceptibility in Plant

Disease," Ann. Rev. Phytopathol., 12:167-179 (1974)). For many years, Southern Corn Leaf Blight was only known as an endemic disease and was not considered to be major economic importance in the United States. But in 1970, it suddenly broke into a severe epidemic that destroyed 15% of the U.S. corn crop and caused losses estimated at more than $1 billion. This serious damage made Southern Corn Leaf Blight one of the most widely known crop diseases in the U.S.

Prior to the outbreak of the disease, only one race of *C. heterostrophus* (race O) was known in the field. In late 1969 when the disease became an epidemic, a new race of the fungus was identified from infected corn leaves collected in severely diseased areas. It was soon designated as race T because of its high virulence on T-cytoplasm corn and the ability to produce a phytotoxin called T-toxin, which specifically affects T-corn. In contrast, race O does not produce T-toxin and is mildly virulent on both T-cytoplasm and N-cytoplasm (normal cytoplasm) corn (Hooker et al., "Physiological Races of *Helminthosporium maydis* and Disease Resistance," Plant Dis. Reptr., 54:1109-1110 (1970); Scheifele, "Cytoplasmically Inherited Susceptibility to Diseases as Related to Cytoplasmically Controlled Pollen Sterility in Maize," 25:110-138 (1970); Smith et al., "Physiologic Races of *Helminthosporium maydis*," 54:819-822 (1970); Yoder et al., "Segregation of Pathogenicity Types and Host-Specific Toxin Production in Progenies of Crosses Between Races T and O of *Helminthosporium maydis* (*Cochliobolus heterostrophus*)," Phytopathology, 65:273-276 (1975); Yoder, "Evaluation of the Role of *Helminthosporium maydis*, Race T Toxin in Southern Corn Leaf Blight, in Tomiyama, eds., Biochemistry and Cytology of Plant Parasite Interaction, New York, New York:Elsevier, pp. 16-24 (1976); Yoder, "Toxins in Pathogenesis," Ann. Rev. Phytopathol., 18:103-129 (1980)). T-cytoplasm stands for Texas male sterile cytoplasm, a unique cytoplasm with a trait for maternally inherited male sterility, characterized by the failure to produce pollen (Levings, "The Texas Cytoplasm of Maize: Cytoplasmic Male Sterility and Disease Susceptibility," Science, 250:942-947 (1990)). T-cytoplasm corn was widely used for hybrid seed production and breeding to avoid hand or mechanical emasculation in the 1950s and the 1960s. It was the coexistence of large acreages of intensively planted T-cytoplasm corn and the

sudden appearance of race T of *C. heterostrophus* that resulted in the epidemic of the disease in 1970. This discovery first opened the door to understanding pathogenesis by *C. heterostrophus*.

Early genetic analysis suggested that both T-toxin production and high virulence on T-cytoplasm corn are controlled by a single genetic locus defined as *Tox1* (Leach et al., "Dominance at the *Tox1* Locus Controlling T-Toxin Production by *Cochliobolus heterostrophus*," Physiol. Plant Pathol., 21:327-333 (1982)). This was demonstrated by crosses between race T and race O in which only parental phenotypes segregated in a 1:1 ratio (Tox^+:Tox^-); all T-toxin producing progeny are highly virulent on T-cytoplasm corn while all T-toxin nonproducing progeny are weakly virulent (Yoder et al., "Segregation of Pathogenicity Types and Host-Specific Toxin Production in Progenies of Crosses Between Races T and O of *Helminthosporium maydis* (*Cochliobolus heterostrophus*), Phytopathology, 65:273-275 (1975); Leach et al., "Dominance at the *Tox1* Locus Controlling T-Toxin Production by *Cochliobolus heterostrophus*," Physiol. Plant Pathol., 21:327-333 (1982)). Further investigation by comparison of electrophoretic karyotypes and chromosome-specific DNA hybridizations indicated that *Tox1* is tightly linked to a reciprocal translocation breakpoint and is associated with as much as a megabase of DNA (mostly highly repeated and A+T-rich) that is missing in race O (Bronson, "Ascospore Abortion in Crosses of *Cochliobolus heterostrophus* Heterozygous for the Virulence Locus *Tox1*," Genome, 30:12-18 (1988); Tzeng et al., "A Restriction Fragment Length Polymorphism Map and Electrophoretic Karyotype of the Fungal Maize Pathogen *Cochliobolus heterostrophus*," Genetics, 130(1):81-96 (1992); Chang et al., "A Reciprocal Translocation and Possible Insertion(s) Tightly Associated with Host-Specific Virulence in *Cochliobolus heterostrophus*," Genome, 39(3):549-557 (1996)). Surprisingly, recent analyses of several Tox^- mutants revealed that *Tox1* is not a single locus but rather two loci, each on a different translocated chromosome (Yoder et al., "Molecular Determinants of the Plant/Fungus Interaction," in Kohmoto, eds., Host-Specific Toxin: Biosynthesis, Receptor and Molecular Biology, Tottori, Japan:Faculty of Agriculture, Tottori Univ., pp. 23-32 (1994); Turgeon et al., "Function and Chromosomal Location of the *Cochliobolus*

heterostrophus Tox1 Locus," Can. J. Bot., 73 (suppl. 1 Sect. E-H):S1071-S1076 (1995)). These two *Tox1* loci have been designated *Tox1A* and *Tox1B* (Yoder et al., "*Cochliobolus spp.* and Their Host-Specific Toxins," in Carroll, eds., The Mycota Vol. 5: Plant Relationships, Part A, Berlin:Springer-Verlag, pp. 145-166 (1997)). Two genes *PKS1* and *DEC1* have been cloned from the two loci respectively; both are required for biosynthesis of T-toxin and are found only in race T isolates of *C. heterostrophus* (Yang, "The Molecular Genetics of T-Toxin Biosynthesis by *Cochliobolus heterostrophus*," Ph.D. Thesis, Cornell University (1995); Yang et al., "A Polyketide Synthase is Required for Fungal Virulence and Production of the Polyketide T-Toxin," Plant Cell, 8(11):2139-2150 (1996); Rose et al., "A Decarboxylase Required for Poloyketide Toxin Production and High Virulence by *Cochliobolus heterostrophus*," 8th Int. Symp. Mol. Plant-Microbe Int., Knoxville, p. J-49 (1996)).

Genetic analysis also suggested that T-toxin is required by *C. heterostrophus* for its high virulence on T-cytoplasm corn. This hypothesis was first tested by the generation of induced T-toxin deficient mutants using different mutagenesis procedures. All mutants with a tight Tox^- phenotype cause disease symptoms that are indistinguishable from those caused by race O when tested on both T and N-cytoplasm corn, suggesting that T-toxin is indeed a virulence factor (Yang et al., 1992; Lu et al., "Tagged Mutations at the *Tox1* Locus of *Cochliobolus heterostrophus* Using Restriction Enzyme-Mediated Integration," Proc. Natl. Acad. Sci. USA, 91:12649-12653 (1994); Rose et al., "A Decarboxylase Required for Polyketide Toxin Production and High Virulence by *Cochliobolus heterostrophus*," 8th Int. Symp. Mol. Plant-Microbe Int., Knoxville, p. J-49 (1996)). This conclusion was firmly supported by the site-specific disruption of the *PKS1* or *DEC1* in the wild type race T genome; disruptants lost the ability to produce T-toxin and caused race O type symptoms on both T-corn and N-corn (Yang et al., "A Polyketide Synthase is Required for Fungal Virulence and Production of the Polyketide T-Toxin," Plant Cell, 8(11):2139-2150 (1996); Rose et al., "A Decarboxylase Required for Polyketide Toxin Production and High Virulence by *Cochliobolus heterostrophus*," 8th Int. Symp. Mol. Plant-Microbe Int., Knoxville, p. J-49 (1996)). These experiments have given a very

clear resolution for the role of T-toxin in pathogenesis. They also implied that pathogenesis by *C. heterostrophus* must involve additional pathogenicity factors because race O which does not produce T-toxin and race T-derived *Tox*$^-$ mutants are effective pathogens on corn.

A number of fungal molecules have been identified as general pathogenicity or virulence factors in several plant pathogenic fungi (Yoder et al., "Molecular-Genetic Evaluation of Fungal Molecules for Roles in Pathogenesis in Plants," J. Genet., 75(3):425-440 (1996)). These include potential penetration factors such as melanin (Guillen et al., "Linkage Among Melanin Biosynthetic Mutations in *Cochliobolus heterostrophus*," Fungal Genet. Newsl., 41:41-42 (1994)), cutinase (Oeser et al., "Pathogenesis by *Cochliobolus heterostrophus* Transformants Expressing a Gene Encoding Cutinase from *Nectria haematococca*," Mol. Plant-Microbe Int., 7:282-288 (1994)) and polygalacturonase and xylanase (Lyngholm et al., "Mutants of *Cochliobolus heterostrophus* Deficient in Extracellular Enzymes," Fungal Genet. Newsl., 42:46-47 (1995)) or possible mechanisms involved in colonization such as phytotoxin detoxification (Schäfer et al., "One Enzyme Makes a Fungal Pathogen, But Not a Saprophyte, Virulent on a New Host Plant," Science, 246:247-249 (1989)) or components of signal transduction pathways (Horwitz et al., "A G Protein Alpha Subunit Gene From the Corn Pathogen *Cochliobolus heterostrophus* is Involved in Two Complex Developmental Pathways: Mating and Appressorium Formation (unpublished) (1997)). Although *C. heterostrophus* is known to produce a nonhost specific toxin called ophiobolin (or cochliobolin), a C_{25} sesterterpenoid compound, which is toxic to many organisms, including plants, bacteria, fungi and nematodes, there is no evidence that ophiobolins are involved in pathogenesis by *C. heterostrophus* or other phytopathogenic fungi. No other pathogenesis-related toxins have been isolated from *C. heterostrophus* so far, but studies on closely related *Cochliobolus* species and other phytopathogenic fungi suggest that pathogenesis by this group of fungi also involves peptide toxins.

Four peptide phytotoxins (victorin, HC-toxin, AM-toxin, and enniatins) have been characterized as pathogenicity or virulence factors. They are

all small cyclic peptides (4-6 residues), containing unusual amino acids or hydroxy acids, and they can be either host specific or non-host specific in terms of plant toxicity. A number of peptide phytotoxins are believed to be synthesized nonribosomally. Early in the 1960s, several biochemists working on the bacterial peptide antibiotics gramicidin and tyrocidine found that these polypeptides can be synthesized in RNAase-treated particle-free extracts of *Bacillus brevis* that are known to produce the same antibiotics; adding protein-synthesis inhibitors to the extracts does not affect this process. This indicated the existence of a peptide biosynthetic system in which ribosomes and mRNAs are not needed. Further studies revealed that in this system, peptides are synthesized on a protein-template and this template itself is a multifunctional enzyme or a complex of several such enzymes, collectively called peptide synthetases, catalyzing the biosynthetic process (Laland et al., "The Protein Thiotemplate Mechanism of Synthesis for the Peptide Antibiotics Produced by *Bacillus Brevis*," Essays in Biochemistry, 7:31-57 (1973); Lipmann, "Bacterial Production of Antibiotic Polypeptides by Thiol-Linked Synthesis on Protein Templates," Adv. Microbiol. Physiol., 21:277-266 (1980)).

Peptide synthetases can catalyze biosynthesis of a variety of peptides. In terms of bioactivity, they can be antibiotics, enzyme inhibitors, plant or animal toxins and immunosuppressants (Stachelhaus et al., "Modular Structure of Peptide Synthetases Revealed by Dissection of the Multifunctional Enzyme GrsA.," Journal of Biological Chemistry, 270(11):6163-6169 (1995)). In terms of chemical structure, they can be either linear (i.e. ACV, the penicillin precursor and gramicidin) or cyclic (most are). The latter can be further classified into three subgroups: 1) The "standard" cyclic peptides (i.e. gramicidin S, tyrocidine, HC-toxin and cyclosporin); 2) cyclic lactones (i.e. destruxin); 3) cyclic depsipeptides (i.e. beauvericin and enniatin). There have been over 300 different carboxy compounds that can be activated by peptide synthetases.

Although the first peptide synthetase, Gramicidin S synthetase, was purified and used for the cell-free synthesis of the peptide early in the 1960s (Tomino et al., "Cell-Free Synthesis of Gramicidin S," Biochem., 6:2552-2560 (1967)), the first bacterial peptide synthetase gene, *tycA*, which encodes the

tyrocidine synthetase 1 in *B. brevis* was not cloned until almost twenty years later (Marahiel et al., "Cloning of the Tyrocidine Synthetase 1 Gene from *Bacillus-brevis* and Its Expression in *Escherichia-coli*, Mol. Gen. Genet., 201(2):1986 (1985)). Since then, more than twenty peptide synthetase genes have been reported for both bacteria and filamentous fungi, but only fourteen have complete nucleotide sequences published. All are larger than 3.3 kb and range between 3.3-19.5 kb for bacterial genes and 9.4-45.8 kb for fungal ones. Interestingly, all fungal peptide synthetase genes reported lack introns, even the cyclosporin A synthetase gene *simA*, which has a 45.8 kb of open reading frame (the largest genomic ORF so far recorded). Although biosynthesis of bacterial peptides differs from that of fungal ones in terms of the number of multifunctional enzymes involved, the genes encoding these enzymes are similar to each other in both function and structure. Comparison of nucleotide sequences reveals one or more highly conserved regions at certain positions in each peptide synthetase gene. These regions formerly called "amino acid activating domains" (Stachelhaus et al., "Modular Structure of Peptide Synthetases Revealed by Dissection of the Multifunctional Enzyme GrsA," Journal of Biological Chemistry, 270(11):6163-6169 (1995)), now called "amino acid activating modules" (Marahiel, "Protein Templates for the Biosynthesis of Peptide Antibiotics," Chem. Biol., 4(8):561-567 (1997)) consist of a set of domains (formerly called "modules") believed to have specific functions such as recognization, activation and thioesterification of individual constituent amino or hydroxy acids, and in some cases methylation and racemation for modification of certain residues before incorporation into the peptide chain (Stachelhaus et al., "Modular Structure of Peptide Synthetases Revealed by Dissection of the Multifunctional Enzyme GrsA," Journal of Biological Chemistry, 270(11):6163-6169 (1995)). The most convincing evidence supporting this assignment is that in most cases, the number of conserved functional units in each gene or gene cluster is equal to the number of amino acids in the respective peptide. This one-for-one match is very clear between three of four fungal peptides and their biosynthetic genes. The total number of modules in three of four bacterial gene clusters also matches the number of amino acids in the respective peptides.

Sequence alignment of amino acid-activating modules reveals strictly conserved sequence motifs that contain active residues for module functions. These motifs are called "core sequences" (Marahiel, "Multidomain Enzymes Involved in Peptide Synthesis," FEBS Lett., 307(1):40-43 (1992)). A minimal amino acid-activating module must contain six core sequences, whose functions (except for core 1) have been proposed based on mutational analysis of several peptide synthetases. Core sequences 1-5 are grouped into an amino acid adenylation domain and core 6 is a thioester formation domain (Figure 1A). All bacterial peptide synthetase genes contain "type I modules" – the minimal amino acid activating modules which were previously called "type I domains" (Stachelhaus et al., "Modular Structure of Peptide Synthetases Revealed by Dissection of the Multifunctional Enzyme GrsA," Journal of Biological Chemistry, 270(11):6163-6169 (1995)). Two fungal genes, *acvA* and *HTS1* also have this modular structure. In addition to the type I module, two fungal genes, *esyn1* and *simA*, contain type II modules, in which an insertion (about 400 amino acids) is found between cores 5 and 6 of a normal type I module. This region contains a motif (VLE/DXGXGXG (SEQ. ID. No. 1)), highly conserved in *S*-adenosyl-methionine (SAM)-dependent methyltransferases, hence, it is referred to as a *N*-methylation domain (Figure 1A). Additional evidence for methyltransferase activity of this module is that the number and position of type II modules in *esyn1* and *simA* exactly match that of *N*-methylated amino acids in ennatin and cyclosporin sequences (Figure 1B).

Although the modular structure described above is highly conserved among most peptide synthetase genes, some variations have been found in the latest cloned peptide synthetase gene *safB*, which is the first gene in the saframycin Mx1 synthetase gene cluster (Pospiech et al., "A New *Myxococcus xanthus* Gene Cluster for the Biosynthesis of the Antibiotic Saframycin Mx1 Encoding a Peptide Synthetase," Microbiology, 141(8):1793-1803 (1995)). *safB* contains two type I amino acid activating modules. One module has all six highly conserved core sequences, but another, believed to activate alanine (the first amino acid in the linear tetrapeptide precursor of saframycin Mx1), lacks core 5 and has a weakly conserved core 1 (Pospiech et al., "Two Multifunctional Peptide

Synthetases and an O-methyltransferase Are Involved in the Biosynthesis of the DNA-Binding Antibiotic and Antitumour Agent Saframycin Mx1 from *Myxococcus xanthus*," Microbiology, 142(4):741-746 (1996)) (Figure 1A). This suggests that some of the motifs in the amino acid adenylation domain are dispensable or not critical for domain function. It also raises the possibility that other variations might be found in yet unknown peptide synthetase genes.

Although *C. heterostrophus* has been a model eukaryotic plant pathogen since the 1970s, most molecular genetic analyses conducted in this system have focused on production of the polyketide T-toxin by race T isolates of the fungus. Solid evidence now indicates that T-toxin is a host-specific virulence factor in Southern Corn Leaf Blight (Yoder et al., "Molecular-Genetic Evaluation of Fungal Molecules for Roles in Pathogenesis in Plants," J. Genet., 75(3):425-440 (1996); Yoder et al., *Cochliobolus spp.* and Their Host-Specific Toxins, Carroll eds., The Mycota Vol. 5: Plant Relationships, Part A, Berlin:Springer-Verlag, pp. 145-166 (1997)). It is clear, however, that *C. heterostrophus* needs additional factors, presumably general factors for pathogenesis to corn plants, since race O, which does not produce T-toxin, can be an effective corn pathogen. Attempts to identify additional general factors required by *C. heterostrophus* for pathogenesis have been unsuccessful. Cloning and characterizing additional *C. heterostrophus* genes that control biosynthesis of novel fungal molecules involved in critical pathogenic processes may be important because it offers potential targets for the design of products that might interfere with the corn plant infection process. The present invention is directed to achieving these objectives.

SUMMARY OF THE INVENTION

The present invention relates to an isolated nucleic acid molecule encoding CPS1 peptide synthetase homologs. The DNA molecule comprises a nucleotide sequence which hybridizes to a DNA molecule having a sequence as set forth in at least one of SEQ. ID. No. 2, SEQ ID No. 41, SEQ ID No. 43, or SEQ ID No. 45.

Another aspect of the present invention relates to a method for identifying inhibitors of a CPS1 protein or polypeptide which involves providing

the CPS1 protein or polypeptide, contacting the protein with potential inhibitor compounds, determining peptide synthetase activity, and selecting compounds which decrease the peptide synthetase activity.

Still another aspect of the present invention relates to a method of imparting disease resistance to a plant by overexpressing a CPS1 protein or polypeptide in the plant cell.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides the structure of amino-acid activating modules identified in peptide synthetase genes (adapted from Stachelhaus and Marahiel, 1995; Pospiech 1995; Marahiel, 1997). Figure 1A shows the domain arrangements in two types of modules. Structural variations in the first module (safB1) of the gene *safB* are also indicated below type I. Figure 1B shows the correlation between module types and the nature of residues in two fungal peptides. Open box: type I module; filled box: type II module. Each peptide sequence is given below.

Figure 2 depicts the recovery of DNA flanking the REMI vector insertion site (arrows) in mutant R.C4.2696. Circled numbers indicate restriction enzyme sites used for recovery of each plasmid. p214B7 contains 4.2 kb flanking DNA (3.4 left; 0.7 right); p214M1 contains 0.1 kb left flank that overlaps with p214B7 and 1.1 kb right flank that overlaps with p214S1, which contains 3.2 kb flanking DNA on the left only.

Figures 3A-B illustrate the extension of the tagged sequence by targeted integration and plasmid rescue. Figure 3A shows a general scheme illustrating chromosome walking strategy (only one direction is indicated). "X" indicates a restriction enzyme site used for recovery of vector (indicated by "T") with flanking genomic DNA ("L" for left flank; "R" for right flank; R' for a fragment from "R" that is subcloned into a subsequent transformation vector). Each integration site is indicated by a vertical arrowhead pointed at the selectable marker on the vector. Overlapped flanking DNA is represented by bars drawn in the same pattern. Figure 3B shows two strategies used for construction of a

targeted integration vector. Left: A genomic DNA fragment (indicated by the letter "a") is subcloned into a vector directly and the subsequent integration occurs by a single crossover ("a" is duplicated). Right: "a" is subcloned into a vector *via* multiple steps (see Figure 6 for an example) that allows linearization with restriction enzyme "Y" and the subsequent integration occurs by a double crossover resulting in a simple insertion.

Figure 4 shows how the targeted integration vector p214SNP was constructed. The sites for restriction enzyme digestion are indicated by arrows. p214SNP was originally designed for a double crossover integration after digested with *Bst*XI and *Bcl*I, but it was found later that the *E. coli* strain (DH5α) used for amplification of the plasmid is Dam^+ which blocks the *Bcl*I site (TGA TCA to TG^mTCA). As an alternative, the plasmid was linearized with *Bst*XI only and used for transformation.

Figure 5 depicts the recovery of genomic DNA flanking the targeted integration vector p214SNP in transformant #118. The vector integrated into the target site by a single crossover that resulted in duplicated *Hin*dIII-*Sac*I fragments (indicated by letter b and b'). The second genomic DNA fragment carried by the vector is also duplicated in the genome (*Nar*I-*Hin*dIII, indicated by letter a and a'). Genomic DNA was digested with *Bgl*II (which does not cut the vector) or *Bcl*I (which cuts once in fragment a' on the vector) as indicated by numbers in a circle. The two recovered plasmids carry sequences that extend the right flank only by 170 and 800 bp respectively beyond the genomic DNA on p214S1.

Figure 6 shows how the targeted integration vector p118BSP was constructed. The sites for restriction enzyme digestion on the vectors are indicated by arrows. p118B14 contains two duplicated *Nar*I-*Hin*dIII-*Sac*I fragments (numbered region, a + b = a' + b') but only one is present in p118BSP. p118BSP was linearized with *Bgl*II and transformed into a wild type C. *heterostrophus* strain.

Figure 7 depicts the recovery of genomic DNA flanking the targeted integration vector p118BSP in transformant #9. The vector integrated into the target site by a double crossover. Genomic DNA was digested with *Pst*I and relegated. The recovered plasmid p9P2 includes the entire pUC18 sequence on p118BSP and 4.6 kb of genomic DNA that contains all of ORF1 (*CPS1*), including the stop codon (TAG) and 3.0 kb of genomic region 3' of the stop codon.

Figure 8 shows how the targeted integration vector p118BCS was constructed. The sites for restriction enzyme digestion are indicated by arrows.

Figure 9 depicts the recovery of genomic DNA flanking the targeted integration vector p118BSP in transformant #12. The vector integrated into the target site by a single crossover that resulted in a duplicated *Ssp*I-*Ssp*I fragment in the transformant genome (indicated by the letter "a"). Genomic DNA was digested with *Hin*dIII (arrows) and ligated. The recovered plasmid p12H6 contains the entire p118BSP sequence and a 2.1 kb genomic DNA (*Sac*I-*Hin*dIII) on the left region that overlaps with the sequence carried on p9P2 (see Figure 7).

Figures 10A-B are photographs which show that the REMI mutant R.C4.2696 grows like wild type in culture (Figure 10A) and produces normal appressoria (Figure 10B). In Figure 10A, plates containing complete medium (CM) were inoculated with a conidia-bearing mycelium plug and incubated at 22°C under warm white light (F40/350BL) (Sylvania Inc., Danvers, MA). The photograph in Figure 10A was taken 6 days after inoculation. Left to right: Mutant R.C4.2696; wild type. In Figure 10B, conidia of mutant (left) or wild type (right) were placed in a drop of water and incubated at 32°C for 6 hrs. No significant difference in percentage of appressorium-forming conidia (arrows) was detected.

Figures 11A-B are photographs which show that the REMI mutant R.C4.2696 produces wild type levels of T-toxin (Figure 11B) but has reduced virulence on T-cytoplasm corn (Figure 11A). In Figure 11A, two week old T-cytoplasm corn plants were inoculated with conidial suspensions of (left to right)

mutant R.C42696, wild type race T, and a Tox+ mutant progeny from a cross between R.C4.2696 and a race O tester, and symptoms (represented by three diseased leaves for each strain) were recorded 5 days after inoculation. Leaves infected by wild type race T collapsed (arrow) but those infected by mutants remained alive. In Figure 11B, a plate containing T-toxin-sensitive E. coli cells was inoculated with agar blocks bearing mycelia of three strains (in the same order as in Figure 11A) and incubated at 32°C overnight (the inoculum at bottom is race O control). No significant differences in T-toxin production (indicated by halos) among the three strains were detected.

Figure 12 is a photograph which displays that the REMI mutant R.C4.2696 produces lesions much smaller than wild type on N-cytoplasm corn. Two week old N-cytoplasm corn plants were inoculated with conidial suspensions at the same concentration and symptoms were recorded 7 days after inoculation. The mutant (right) produces the same number of lesions as wild type (left) but the size of lesions was dramatically reduced compared to wild type (arrows).

Figures 13A-B show that the REMI mutant R.C4.2696 has a 60% reduction of virulence compared to wild type. In Figure 13A, lengths of 100 typical lesions from corn leaves inoculated with wild type race O and a mutant progeny R45 (*Tox*$^-$, *hygB*^{R}) carrying the R.C4.2696 mutation were measured 7 days after inoculation and values plotted. Figure 13B shows the statistical analysis that 86% of the mutant lesions are less than 4 mm in length (average size 3.5 mm), 60% reduced compared to that of wild type (8.5 mm).

Figures 14A-B illustrate the genetic analysis showing that a tagged, single site mutation is responsible for the mutant phenotype. Figure 14A is a photograph of a plant assay where N-cytoplasm corn was inoculated with parents and progeny indicated in Figure 14B (a complete tetrad from the second cross is shown here; random spore assay for the two crosses gave the same results). Figure 14B is a table which shows progeny segregation data. In both crosses, progeny segregated 1 : 1 for parental type only and all hygromycin B resistant progeny showed the same small lesion phenotype (represented by progeny 1-4 in Figure 14A) as the mutant parent (parent 1 in Figure 14A); all hygromycin B

sensitive progeny show the large lesion phenotype (represented by progeny 5-6 in Figure 14A) of the wild type parent (parent 2 in Figure 14A).

Figure 15 is a restriction map of the cloned sequences surrounding the tagged site. A 11.3 kb genomic region (thick line) was cloned and completely sequenced. The original REMI insertion point in the mutant R.C4.2696 is indicated by a vertical arrow. The asterisks indicate two targeted integration sites in the wild type genome. Two open reading frames (in opposite directions), ORF1 (*CPS1*, 5.4 kb) and ORF2 (*TES1*, 1.1 kb) are indicated by open boxes below the map (the positions of putative introns are indicated by vertical bars). Locations of seven overlapping plasmid clones used for sequencing are indicated by thin lines on the top of the map (filled triangles represent the vector sequence in each clone). Sequencing strategy is indicated by arrows above each clone line.

Figure 16 graphically depicts the G+C content of the cloned sequences surrounding the tagged site. The 11.3 kb of genomic DNA sequence (Figure 15) was examined as 200 bp fragments and G+C content of each fragment was calculated and plotted. Regions A and C have normal G+C content; regions B and D show high G+C content. The overall percentage of G+C in each region is indicated by underlined numbers. The positions of ORF1 (*CPS1*) and ORF2 (*TES1*) are indicated by arrows and the percentage of G+C of each ORF is given below (in parentheses).

Figure 17 displays the nucleotide sequence of *CPS1*. 5,725 base pairs from the 11.3 kb sequenced region (Figure 15) are shown. The deduced amino acid sequence of CPS1 protein is given below the DNA sequence. The position of start codon ATG (bold and underlined) is designated +1 and the open reading frame stops at position 5381 (TAG, in bold and underlined). Five putative "CAAT" boxes (bold and indicated by asterisks) are found at the positions -36, -58, -67, -172 and -309. Three putative introns (in lowercase with 5' and 3' splice sequences in bold; branch sites are underlined) are located at positions 2070-2114, 3542-3592 and 4197-4249. Conserved core sequences are shaded. The GXSXG motif is boxed. A putative polyadenylation signal is located at position 5604 (bold and overlined).

Figure 18 displays the nucleotide sequence of *TES1*. 1,901 base pairs 5' of *CPS1* are shown. The deduced amino acid sequence of the TES1 protein is given below the DNA sequence. The CT motif (bold and indicated by asterisks) is found at position -46. A "AT"-rich region (bold and italicized) is found at position -12. The position of ATG start codon (bold and underlined) is designated +1 and the open reading frame (transcribed in the opposite direction from *CPS1*) stops at position 1153 (TGA, in bold and underlined). One putative intron (in lowercase, border sequences are bold; branch sites are underlined) is located at position 518-566. A putative polyadenylation signal is located at position 1345 (bold and overlined). The putative active site (VHS) is shaded.

Figures 19A-C are schematic representations which show the characterization of modular structure of *CPS1*. Peptide synthetase and thioesterase are indicated by open boxes; shaded boxes inside indicate functional domains and modules; vertical bars in the shaded boxes indicate highly conserved core sequences. Figure 19A illustrates the general structure of bacterial and fungal peptide synthetases (adapted from Marahiel 1997, which is herein incorporated by reference). A peptide synthetase gene cluster is shown on the top. There can be one or more amino acid activating module (cyclosporine synthetase has 11) in each protein; some peptide synthetases have thioesterase domains (TE), which can be either integrated into modules or encoded by a separate gene. Each synthetase can have type I, type II or both modules. A type I (minimal) module is enlarged to show organization of core sequences and domains. Some peptide synthetases also have condensation or epimerization domains. Figure 19B illustrates the organization of saframycin Mx1 synthetase containing 4 amino acid activating modules (Pospiech et. al., "Two Multifunctional Peptide Synthetases and an O-methyltransferase are Involved in the Biosynthesis of the DNA-Binding Antibiotic and Antitumour Agent Saframycin Mx1 from *Myxococcus xanthus*," Microbiology, 142(4):741-746 (1996)). SafB1 from the first module is enlarged. Core sequences 1 and 5 in safB1 are weakly conserved (indicated by dashed vertical bars). The remaining domains are typical of type I as shown in Figure 19A. SafC is a putative O-methyltransferase. Figure 19C illustrates the organization of *CPS1*. Sequence analysis revealed two amino acid activating

modules (CPS1A and CPS1B), both of which have high similarity to safB1 except that core 2 is weakly conserved. A thioesterase domain is found at the *C*-terminal region of CPS1B. Three vertical arrows indicate the positions of targeted gene disruptions in the wild type genome that yielded the mutant phenotype. TES1 is a thioesterase encoded by a separate gene (*TES1*).

Figures 20A-F are comparative alignments of core amino acid sequences in CPS1A and CPS1B with those of other peptide synthetases. Figures 20A-E are comparative amino acid sequence alignments of amino-acid-activating domains (cores 1-5); Figure 20F is a comparative amino acid sequence alignment of the thioester formation domain (core 6). In each subfigure, the first column shows the names of peptide synthetases; the second indicates the position of the first residue aligned in the original amino acid sequence of each protein; the last column on the right indicates the number of amino acids between two cores (Figures 20A-E, in parentheses) or the distance between two adjacent amino-acid-activating modules (Figure 20F, in parentheses). There is an extra column on the right in Figure 20F, showing the total number (underlined) of residues in each amino-acid-activating module in which the aligned core sequence is located. The consensus of each core sequence is on the top, which includes identical or similar residues found in all peptide synthetases or with only a few exceptions (active site is also indicated by asterisks). SafB1: the first module in saframycin Mx1 synthetase B of *Myxococcus xanthus* (Genbank accession U24657); GrsA: gramicidin S synthetase A of *Bacillus brevis* (SWISS PROT accession P14687); HTS1A and HTS1B: the first two modules in HC-toxin synthetase of *Cochliobolus carbonum* (Q01886); EsynA and EsynB: two modules in enniatin synthetase of *Fusarium scirpi* (EMBL accession Z18755); ACVA and ACVB: the first two modules in ACV synthetase of *Aspergillus nidulans* (SWISS PROT P19787); CsynA and CsynB: the first two modules in cyclosporine synthetase of *Tolypocladium nivenm* (EMBL Z28383).

Figure 21 is a comparative alignment of amino acid sequences of active sites of thioesterase domains (TE) in CPS1 with those of other peptide synthetases. ACV: ACV synthetase (Swiss-PROT accession P19787); GrsB: gramicidin S synthetase B (P14688); GrsT: the thioesterase encoded by *grsT*

(P14686) in gramicidin S synthetase gene cluster; SrfA: surfactin synthetase A-3 (Q08787); TycC: tyrocidine synthetase C (Genbank accession AF004853); TycF: the thioesterase encoded by *tycF* (AF004853) in the tyrocidine synthetase gene cluster. The highly conserved residues (GXSXG) are indicated by asterisks. The number on the left of each amino acid sequence indicates the original position of the first residue; the number on the right (in parentheses) indicates the distance between the last residue shown to the end of each protein.

Figure 22 is a comparative alignment of the amino acid sequence of the *TES1* protein (CH-TES1) with that of other type II thioesterases. HS-TEII: *Homo sapiens* thioesterase II (EMBL accession X86032); EC-TESB: *E. coli* acyl-coA thioesterase II (Genbank accession M63308); MT-TESB: *Mycobactrium tuberculosis* homolog to *E. coli.* acyl-coA thioesterase II (EMBL Z95387). The identical residues in all four proteins are in bold. The putative active site VHS motif is indicated by asterisks. The numbers on the right column indicate the original position of the last residue of the line in each protein sequence. The entire protein sequence of each TES was aligned using the Jotun Hein Method. Amino acids corresponding to the positions 142-171, 236-265 and 356-367 in CH-TES1 have no significant similarity among the four proteins.

Figures 23A-B are photographs displaying a plant assay (Figure 23A) and a gel blot (Figure 23B). Targeted gene disruption suggests that *CPS1* is involved in fungal pathogenesis. In Figure 23A, N-cytoplasm corn was inoculated with (left to right): The REMI mutant R.C4.2696; wild type race T; wild type race O and five disruptants obtained using the linearized p214B7 as vector. All disruptants give the small lesions similar to the original mutant. In Figure 23B, total genomic DNA was digested with *BglII* and probed with both 5' and 3' end flanking DNA fragments carried on p214B7 separately (strain order is the same as above). A single band (4.2 kb) is present in both wild type race T and race O but replaced by a 9.3 kb band (increased by the size of the vector, 5.1 kb) in all strains that showed the mutant phenotype in Figure 23A.

Figure 24 displays a gel blot analysis showing targeted integration of the chromosome walking vector p214SNP into the wild type genome. Lane 1:

wild type race O. Lanes 2-8: seven transformants obtained by transforming wild type race O (strain C5) with p214SNP. Genomic DNAs were digested with *Bgl*II and probed with 3.2 kb flanking genomic DNA fragment cloned on p214S1 which gives three bands in wild type (4.2, 2.1 and 0.5 kb, indicated by arrows on the left). The targeted site is in the 2.1 kb fragment which is missing in all disruptants (indicated by an arrow on the right). Other two bands are intact as predicted. All disruptants showed the same mutant phenotype in the plant assay as shown in Figure 23. Five disruptants (lanes 2, 3, 5, 7 and 8) resulted from a single crossover integration. Integration in two other disruptants (lanes 4 and 6) has not been determined. One of the disruptants (#118, lane 2) was used to recover the plasmid p118B14 and p118BC4.

Figure 25 displays a gel blot analysis showing targeted integration of the chromosome walking vector p118BSP into the wild type genome. Lane 1: wild type race O; Lane 2: wild type race T; Lanes 3, 4 and 5: three transformants obtained by transforming wild type race O (strain C5) with p118BSP. Genomic DNAs were digested with *Pst* I (which cuts pUCATPH) and probed with the 3.2 kb flanking genomic DNA fragment cloned on p214S1 which gives two bands in wild type (6.6 and 2.5 kb, indicated by arrows on the left). The targeted site is in the 6.6 kb region which is missing in all three disruptants (indicated by arrows on the right). The 2.5 kb band is intact as predicted. The third band (part of vector plus 4.6 kb genomic DNA) does not hybridize to the probe. All disruptants showed the same mutant phenotype in the plant assay as shown in Figure 23. Two disruptants (lanes 3 and 4) resulted from a double crossover integration. Integration in the third disruptant (lane 5) has not been determined. One of the disruptants (#9, lane 3) was used to recover the plasmid p9P2.

Figure 26 displays a gel blot showing the detection of *CPS1* homologs in *C. victoriae* and *C. carbonum*. Genomic DNAs were digested with *Bgl*II and probed with the 3.4 kb *CPS1* fragment cloned on p214B7 (Figure 2) which includes most of the 4.2 kb *Bgl*II fragment of *CPS1* (Figure 15). Lanes 1 and 2: *C. heterostrophus* race T (C4) and race O (C5), both of which hybridized to the 4.2 kb fragment. Lanes 3: *C. victoriae* (Hvw). Lanes 4, 5 and 6: *C. carbonum* race 1 (26R13), race 2 (YugY) and race 3 (BZ1209). Note that both *C. victoriae*

and *C. carbonum* (three races) hybridized to a 5.0 kb fragment and the hybridization signals of all three species showed about the same intensity.

Figure 27 is a high-performance liquid chromatography (HPLC) profile of culture extracts from wild type *C. victoriae* (HvW, top left) and three transformants (Tx7, Tx2 and Tx9). The major peak for victorin C is indicated by arrowheads. No significant differences in victorin production were detected between the wild type and transformants. Three other transformants (Tx4, Tx5 and Tx8) gave the same results.

Figure 28A-B are photographs which display a plant assay (Figure 28A) and a gel blot (Figure 28B), showing targeted disruption of the *CPS1* homolog in *C. victoriae*. Figure 28A shows *C. victoriae* transformants (Tx) with reduced or wild type pathogenicity to susceptible oats. Oat seeds were inoculated with conidial suspensions of (left to right) wild type, Tx7, Tx2 and Tx9 (the last pot on the right is the uninoculated control). Two transformants (Tx7 and Tx2) showed dramatically reduced pathogenicity as indicated by the substantial growth of the oat plants. These two transformants resulted from a homologous integration (lanes 2 and 3 in Figure 28B). Tx9, which killed all oat plants as wild type, resulted from a ectopic integration (lane 4 in Figure 28B). All three transformants produced wild type level of victorin as determined by HPLC analysis (Figure 27). Figure 28B shows disruption of *CPS1* homolog in the wild type genome. Genomic DNAs were digested with *Bgl*II and probed with the 3.2 kb *CPS1* fragment (*Kpn*I-*Sac*I) cloned on p214S1 (Figure 2) which hybridized to two fragments (4.2 and 2.2 kb, see Figure 15) in *C. heterostrophus* (lane 1), but to three fragments (5.0, 1.8 and 0.2 kb) in the *C. victoriae* (lane 2). In Tx7 (lane 3) and Tx2 (lane 4), one or two of the wild type fragments was replaced by a larger fragment (8.0 kb for Tx7 and 9.4 kb for Tx2) containing the transforming vector. The actual size increase by the vector integration can not be predicted because the presence of polymorphic bands in *C. victoriae* genome and the presence of duplicated *CPS1* fragments on the transforming vector (Figure 5). In Tx9 (lane 5), which caused wild type symptoms shown in Figure 28A, all three wild type bands (indicated by arrows) are intact, confirming an ectopic integration.

Figure 29 shows the REMI vector pUCATPH. This vector was constructed by insertion of a 2.4 kb *Sal*I fragment containing the selectable marker cassette (the largest arrow) from pDH25 (Cullen et al., 1987) into the *Sal*I site of polylinker site of pUC18 (between *lacI* and *lacZ*). Only six-base-pair restriction enzyme sites are shown on the map. Italicized sites are unique; three of them, *Hin*dIII, *Kpn*I, and *Sac*I (bold) have been used for REMI transformation. Non-cutting enzymes are listed below the map (enzymes that recognize six-base-pair sites are underlined). *amp*, Ampicillin resistance gene; *hygB*, hygromycin B resistance gene; *PtrpC*, *A. nidulans trpC* promoter; *TtrpC*, *A. nidulans trpC* terminator; *ori*, *Escherichia coli* origin of replication.

Figures 30A-30C are photographs of DNA gel blots showing DNA-DNA hybridization of *ChCPS1* to other fungal genera and species. In Figure 30A, the gel was loaded with *Cochliobolus* species (lanes 1-17) as follows: *C. heterostrophs* race T, race O; *C. carbonum* race 1, race 2; *C. victoriae* isolates FI3, HvW; *C. bicolor, C. dactyloctenii, C. chloridis, C. homomorphus, C. intermedius, C. melinidis, C. melinidis, C. peregianensis, C. perotidis, C. ravenelii* and *C. sativus*. Figure 30B is a photograph of a DNA gel blot from a gel loaded with other *Ascomycete genera* (lanes 1-14) as follows: *C. carbonum* race1 (control), *Setosphaeria rostrata, Stemphyllium spp., Pyrenophora tritici repentis, Bipolaris sacchari, Alternaria spp., A. solani, Nectria haematococca, Fusarium oxysporum, Glomerella spp. Magnaporthe grisea, F. moniliforme, F. moniliforme* (repeat) and *A. solani* (repeat). Figure 30C is is a photograph of a DNA gel blot from a gel comparing *Candida albicans* to *C. heterostrophus* and closely related species (lanes 1-7): *C. heterostrophs* race T, *Bipolaris sacchari, Setosphaeria rostrata, Stemphyllium spp., Pyrenophora tritici repentis, Alternaria spp.* and *Candida albicans*(arrowhead). Genomic DNAs were digested with *Hind*III (A, lanes 1-17; B, lanes 1-11; C, lanes 1-7), *Xho*I (B, lanes 12 and 14) or *Bgl*II (B, lane 13) and probed with the 3.2 kb fragment of *CPS1* from p214S1 (Fig. 2) at high stringency. Weak signals in lanes 3 and 17 (panel A) are due to insufficient DNA loading (confirmed by a repeat experiment).

Figure 31A is a structural comparison of the four *CPS1* homologs to *ChCPS1*. ORFs are indicated by the open boxes; shaded boxes inside indicate

functional domains; vertical bars indicate conserved motif sequences found in nonribosomal peptide synthetases (NRPS) as defined by Stachelhaus and Marahiel (Stachelhaus and Marahiel, 1995; Marahiel, 1997) (dashed bars indicate weak conservation). The black bulbs indicate the position of putative introns. Cores 1-5: adenylation; core 6: thioaltion; TE: thioesterase. The distance between core sequences is not drawn in exact scale. The name of proteins is on the left of the ORF boxes and the number of amino acids on the right. The unidentified regions of *AsCPS1* and *PtCPS1* are indicated by dash-lined boxes. The similarity to *ChCPS1*(in the overlapping region only, see text for details) is given in the parentheses under the protein names in the order: nucleotide identity/ amino acid identity/ amino acid similarity. The positions of the ChCPS1 amino acid 1040 is indicated by the open arrow; the positions 511 and 1269 (to the first and the last amino acids of AsCPS1 and PtCPS1) are indicated by filled triangles. Figure 31B is an amino acid alignment of the four CPS1 homologs to ChCPS1. 530 amino acids aligned to the amino acids 511-1040 of ChCPS1(shown in A) are shown. The identical residues are in uppercase and the similar residues in lowercase. Consensus of sequences similar to the typical NRPS signature motifs is underlined. The putative cyclization domain motif "DXXXXD/ EXXS/ A" is underlined.

Figure 32 is the nucleotide sequence of *FgCPS1*. 6,003 base pairs cloned using the plasmid rescue procedure are shown. The amino acid sequence of FgCPS1 protein is given below the DNA sequence. The position of the start codon ATG (bold and underlined) is designated +1 and the open reading frame stops at position 5123 (TGA, bold and underlined). A "CT" motif (italicized and underlined) and two putative "CAAT" boxes (bold with asterisks) are found at positions –30, –204 and -302. A putative intron (in lowercase with 5' and 3' splice sequences in bold; branch sites underlined) is located at positions 4245-4290. Conserved core sequences are shaded and the putative cyclization domain motif "DXXXXEXXA" (position 2323-2346) is underlined. A putative polyadenylation signal "AATAA" at position 5188 is bold and overlined.

Figure 33 shows the nucleotide sequence of *AsCPS1*. 2,369 base pairs amplified by PCR are shown. The amino acid sequence of AsCPS1 protein is

given below the DNA sequence. The sequence is not complete. Two putative introns (in lowercase with 5' and 3' splice sequences in bold; branch sites underlined) are located at positions 540-584 and 2012-2059. Conserved core sequences are shaded and the putative cyclization domain motif "DXXXXDXXS" (positions 694-720) is underlined. The PCR primer binding sites at the 5' and 3' end are underlined.

Figure 34 shows the nucleotide sequence of *PtCPS1*. 2,320 base pairs amplified by PCR are shown. The amino acid sequence of PtCPS1 protein is given below the DNA sequence. The sequence is not complete. A putative intron (in lowercase with 5' and 3' splice sequences in bold; branch sites underlined) is located at positions 540-583. Conserved core sequences are shaded and the putative cyclization domain motif "DXXXXDXXS" (positions 693-719) is underlined. The PCR primer binding sites at the 5' and 3' end are underlined.

Figure 35A is a photograph of a DNA gel blot showing that the 2.2 kb wild type band (arrowhead) is disrupted in homologous transformants TxFgC8-4, -10, -11 and -5 but is intact in the ectopic transformant TxFgC8-H1, -H2 and –B1 [generated using *HindIII* (H) or *BglII*(B)-digested pFgC8-hygB]. Genomic DNAs were digested with *Cla*I and probed with pFgC8 which carries a 1.0 kb *FgCPS1* fragment. Figure 35B is a photograph of plants from a virulence assay showing *F. graminearum* transformants (Tx) with reduced or wild type virulence to wheat. Wheat heads were inoculated with conidial suspensions (10^4/ml) of (left to right) wild type, TxFgC8-4, -10, -H1, and -11 or with water only. Photograph was taken 7 days after inoculation. Note that most spikelets of wheat heads inoculated with homologous transformants looked "healthy" in contrast to those inoculated with ectopic transformants that were completely "bleached" (indistinguishable from wild type).

DETAILED DESCRIPTION OF THE INVENTION

The present invention generally relates to an isolated DNA molecule from a plant pathogen encoding a CPS1 peptide synthetase. In one embodiment, the DNA molecule has a nucleotide sequence which hybridizes to a DNA molecule having a sequence corresponding to SEQ. ID. No. 2 as follows:

TGCCTGCGCC TGTGCTTGTG CCTGTGGAAT GTCGCGGCCC GCTGCTGCAT AGCCTATCTG 60

TACATACAAC ACCATCCCAT CCCGCTTCAC CTGCCTTGCC TCCCTCCTCG TGCCACACAT 120

CCGCCGCCCA CAACACCATG GCTGCGACCA ACCCCGAGCT GCAGGCCAAA CTGCAGGAGC 180

TGGACCACGA GCTCGAGGAG GGCGATATTA CACAAAAAGG GTCCGTACTG CTGCACCACC 240

ACCGCCATCC GCCTCTCTGC GTGCGCTAAT CAGTCGCATA GCTATGAAAA ACGTCGCACC 300

GTGCTGCTGT CGCAGTATCT AGGGCCTGAC TTTGCTGCCC AGTTGCAGGC CGACCTGAAC 360

CAGCAGAACC CACCCCAACC ATCCAGTGAG GGCTCTCGCT CCCGCACCGC ATCCTTTGCT 420

ATTCCGTCCG GTCCGAGTCC ATCACNGCGA CCACAACCCC CACATATCCA GCTCCCCCGC 480

CCCGACTCAT ACCATGACGC TTCCGCACAG GGCCAATTGG GCGCACCCAT GCCATATGCG 540

AACGCCTCCG CCGCTGCCTC GGGGGGCTCG CAGTACATGG CATACCCGCC CAGCCAAGTC 600

GGCCGTTTTC AAGAGAAGCA GCTGGGCCTG CGTACAAATT CGCTCCAGCG CAATTCCTCA 660

CAGCTGTCGC AAGGAAGCGA GACGTTCATT CCACGGCCTC AAACGCCTGA ATACAACCAC 720

TCGCGCGAGC CCACCATGAT GGGCAACTAC GCCTTCAATC CAGACAATCA GCAAAGTTAT 780

GATGGCCAAT TTGGCTCTCC GGGAGAGGCC AGTCGAAGGA GCACCATGCT CGAGGTAAAC 840

CAGGGTTATT TTTCCGACTT CACAGGCCAG CAGATGCAAG ACAATCGCGA CTCGTATGGG 900

GGACCCAACC GCTACTCGTC GGGAGATGCC TTTTCTCCTA CCGCCGCGAT TCCACCTCCC 960

ATGATGAACC CCAACGATCT CCCCTTGGGC GCTGCTGAAA CCATGATGCC GCTAGAGCCC 1020

CGCGATCTGC CTTTTGACGT TTACGACCCT CACAACCCCA ATGTCAAAAT GTCAAAGTTT 1080

GACAACATTG GCGCTGTCTT GCGTCACCGA AGTCGCACAC AGCCAAGGAC GACTGCCTTC 1140

TGGGTCCTTG ACGCAAAAGG CAAAGAGACG GCGTCCATCA CCTGGGAAAA GGTGGCTAGT 1200

CGCGCGGAAA AGGTGGCCAA AGTGATTCGG GACAAGAGCA ACCTCTATCG AGGCGACCGT 1260

GTGGCATTAG TGTACAGGGA TACAGAAATC ATTGATTTTG TCGTGGCGTT GATGGGCTGC 1320

TTCATTGCGG GCGTTGTAGC GGTACCCATC AATAGCGTCG ACGACTACCA GAAACTCATT 1380

CTTCTCCTAA CGACAACTCA AGCTCATCTC GCATTGACCA CAGACAACAA TCTCAAGGCC 1440

TTTCATCGTG ACATTAGTCA GAACCGTCTG AAATGGCCGA GTGGGGTAGA GTGGTGGAAG 1500

ACGAACGAGT TTGGCAGCCA CCACCCCAAG AAACATGACG ATACTCCAGC TTTGCAAGTA 1560

CCAGAGGTTG CCTATATTGA GTTCTCGCGT GCACCTACTG GTGACCTTCG CGGTGTGGTG 1620

CTTAGTCACC GGACTATTAT GCACCAAATG GCCTGCATCA GTGCCATGAT TAGCACGATA 1680

CCCACCAACG CTCAGAGCCA AGACACGTTC AGCACTAGCC TACGGGATGC AGAGGGAAAG 1740

TTCGTTGCTC CAGCACCGTC CAGAAACCCC ACAGAAGTGA TCCTCACGTA CCTCGACCCG 1800

CGCGAAAGCG CTGGTCTCAT TCTCAGTGTC TTGTTTGCAG TTTATGGAGG CCACACCACC 1860

GTATGGCTCG AGACAGCGAC CATGGAAACC CCGGGTCTAT ATGCACATCT CATCACCAAA 1920

TACAAGTCCA ACATACTGCT AGCGGATTAC CCAGGCCTCA AGCGCGCTGC ATACAACTAC 1980

CAACAGGATC CAATGGCTAC AAGAAACTTC AAGAAAAACA CAGAACCCAA CTTCGCCTCC 2040

GTGAAGATCT GTCTGATTGA CACGCTTACC GTCGACTGTG AATTTCACGA AATTCTCGGA 2100

GATCGATATT TCAGGCCACT GCGAAACCCT AGAGCGCGAG AACTGATCGC GCCAATGCTC 2160

TGCTTGCCAG AACATGGTGG AATGATAATA TCTGTACGCG ACTGGCTAGG TGGAGAGGAG 2220

CGCATGGGCT GCCCGCTAAG CATAGCAGTA GAAGAGTCAG ATAATGATGA AGATGATACA 2280

GAGGATAAGT ATGCAGCGGC AAATGGCTAC TCCAGTCTTA TTGGTGGTGG CACTACAAAG 2340

AACAAAAAGG AGAAGAAGAA GAAAGGCCCG ACAGAGCTTA CAGAAATCTT GCTGGACAAG 2400

GAAGCTCTGA AGATGAACGA AGTCATTGTT CTGGCCATTG GAGAAGAAGC AAGCAAGCGG 2460

GCAAACGAGC CCGGCACCAT GCGAGTCGGT GCCTTTGGAT ACCCCATACC GGATGCGACA 2520

CTAGCTATTG TAGACCCTGA GACAAGTCTT CTATGTTCAC CATACTCGAT AGGCGAGATC 2580

TGGGTAGATT CGCCTTCACT CTCTGGTGGC TTCTGGCAGC TGCAGAAGCA TACAGAGACC 2640

ATTTTCCATG CTCGACCATA CCGTTTCGTT GANGGTAGCC CTACGCCACA GTTGCTTGAA 2700

CTCGAGTTTC TGCGTACTGG ACTCCTCGGC TTTGTTGTAG ANGGAAAAAT ATTTGTCCTT 2760

GGACTGTACG AAGATCGCAT CAGACAGCGT GTTGAATGGG TAGAAAATGG TCAGCTTGAA 2820

GCCGAGCATC GATACTTTTT TGTGCAGCAC CTGGTCACAA GCATTATGAA GGCCGTGCCA 2880

AAAATTTACG ACTGGTAAGT GAGCTGCCAA CAGAGCAAGG ACTGTCTAAC GTGTCATAGC 2940

TCGTCGTTTG ATTCTTATGT AAATGGTGAA TACCTGCCAA TCATTCTCAT CGAGACGCAG 3000

GCCGCATCGA CTGCGCCCAC AAACCCAGGT GGACCACCAC AACAATTGGA TATACCATTT 3060

TTGGATTCAC TATCTGAGAG GTGCATGGAG GTCCTTTACC AAGAGCATCA TTTACGGGTA 3120

TACTGCGTGA TGATTACAGC ACCTAATACA CTTCCACGAG TCATCAAGAA CGGACGGCGA 3180

GAAATTGGCA ATATGCTGTG TANGAGANAG TTTGACAATG GCTCTCTGCC CTGTGTNCAC 3240

GTNAAGTTTG GCATTGAGCG ATCAGTGCAG AACATTGCGC TCGGTGACGA TCCCGCTGGC 3300

GGCATGTGGT CATTTGAGGC ATCAATGGCA CGTCAGCAAT TCTTGATGCT CCAAGACAAG 3360

CAATACTCTG GTGTCGATCA TCGCGAAGTC GTCATTGACG ACAGGACATC GACTCCACTC 3420

AATCAGTTCT CGAATATCCA CGACCTGATG CAATGGCGTG TATCTCGGCA GGCCGAGGAA 3480

CTTGCTTACT GCACTGTCGA CGGTCGAGGA AAAGAGGGCA AAGGCGTCAA TTGGAAGAAG 3540

TTTGATCAAA AGGTTGCGGG CGTAGCAATG TACCTCAAGA ACAAGGTCAA GGTCCAGGCC 3600

GGCGATCATC TCCTTCTGAT GTACACGCAT TCAGAAGAAT TTGTTTATGC TGTTCATGCA 3660

TGTTTTGTGC TTGGAGCTGT TTGCATACCA ATGGCGCCAA TTGATCAGAA CCGGTTGAAT 3720

GAGGATGCGC CGGCCTTGCT GCATATCCTT GCAGATTTCA AGGTCAAAGC CATTCTTGTC 3780

AACGCTGACG TTGACCATCT GATGAAGATC AAGCAAGTAT CGCAGCACAT CAAACAATCG 3840

GCCGCTATCC TCAAGATCAG TGTGCCAAAC ACATACAGCA CAACAAAGCC GCCAAAGCAA 3900

TCCAGTGGCT GCCGCGACCT CAAGCTTACA ATTCGACCGG CATGGATTCA GGCGGGTTTC 3960

CCAGTGCTAG TCTGGACATA CTGGACGCCC GATCAACGTC GTATCGCAGT TCAGCTGGGC 4020

CATAGCCAAA TCATGGCACT GTGCAAGGTC CAAAAAGAAA CATGCCAAAT GACAAGTACA 4080

CGACCAGTCC TTGGTTGTGT CCGGAGCACG ATAGGACTTG GTTTCCTTCA CACTTGTCTC 4140

ATGGGAATCT TCCTTGCCGC ACCCACATAC CTGGTGTCAC CTGTTGACTT TGCACAAAAC 4200

CCTAATATTC TGTTCCAAAC GCTTTCGCGG TACAAGATCA AGGATGCATA TGCAACGAGT 4260

CAAATGTTGG ACCACGCCAT CGCACGCGGA GCTGGTAAGA GTATGGCTCT GCACGAGCTG 4320

AAGAATCTCA TGATTGCGAC TGATGGAAGA CCACGCGTTG ATGTTTGTAA GTGAACATTT 4380

GTATGAGAGG ACTTTCATGA TTGCTAACTC AATGCAGACC AAAGAGTGCG TGTGCACTTT 4440

GCGCCAGCCA ACTTAGACCC AACCGCAATC AACACTGTCT ACTCACATGT ATTGAACCCA 4500

ATGGTAGCAT CACGATCATA CATGTGTATT GAGCCAGTCG AGCTCCATCT CGATGTGCAT 4560

GCTCTGCGAC GCGGCCTCGT CATGCCCGTT GACCCTGACA CAGAGCCCAA CGCTTTGCTC 4620

GTCCAAGACT CGGGCATGGT GCCAGTGAGC ACGCAAATAT CCATTGTCAA CCCAGAGACC 4680

AACCAACTGT GCTTGAACGG CGAGTACGGC GAGATCTGGG TGCAGTCCGA GGCGAATGCT 4740

TATAGCTTCT ACATGTCGAA AGAGCGCTTG GATGCAGAAC GCTTCAATGG GAGGACGATT 4800

GACGGAGACC CAAATGTGCG ATATGTTCGT ACAGGCGATT TAGGATTTTT GCACAGCGTG 4860

ACACGGCCCA TTGGACCCAA CGGTGCACCT GTTGATATGC AGGTGCTTTT CGTGCTTGGA 4920

AGCATAGGTG ACACTTTTGA AGTCAACGGA CTGAACCATT TCTCTATGGA CATTGAGCAG 4980

TCTGTTGAAC GTTGTCACCG GAATATTGTC CCTGGAGGCT GGTACGTTTC TTCGATTCGC 5040

TGTTATTTAG TAAATACTTA CTAACACTCT ACAGTGCTGT TTTCCAGGCA GGTGGGCTTG 5100

TTGTTGTCGT TGTGGAAATC TTCCGACGCA ACTTCCTCGC AAGCATGGTG CCTGTGATTG 5160

TCAATGCAAT TTTGAACGAG CATCAGCTGG TCATTGACAT TGTCTCGTTT GTGCAAAAGG 5220

GCGACTTCCA CCGGTCTCGT CTGGGCGAGA AGCAACGCGG AAAGATTCTT GCAGGATGGG 5280

TCACACGGAA GATGCGCACA ATAGCCCAGT ACAGTATACG GGATCCTAAT GGACAGGATT 5340

CCCAGATGAT CACGGAAGAG CCTGGTCCAC GGGCTAGATG ACTGGAAGTA TGCTTGGGCG 5400

AATGGGCGGC CCAGCCAGTA TCAAGGCCGG GTCGACAAGA GCACCGAGTC TAATGGGCAT 5460

GACAGCGACT ATGAATAATC TATCCCTTAC ACAGCAGCAA CAGCAGCAAT ACCAACAGCC 5520

GGGTATGTAT GCTCAACAGC AAGGCATGCA CCCCCAGCAA CAACACCAAT TTAGCATGTC 5580

CAACACGCCA CCACAAGGTC CACCCCAAGG CGTAGAACTA CATGATCCTA GCGACCGCAC 5640

ACCAACAGAC AACCGGCACT CTTTCCTTGC CGACCCGCGT ATGCAGAACC AGGGCCAAAT 5700

GAACGAGACG GGCGCCTACG AACCCATGAA CTATCAAAAC GCGTATCATC CGCATCAACA 5760

ACAATACGAA TCTGAAGACG GGGGGAGCAG ACTCAGCGGC CCCGTGCCAG ACGTGCTGCG 5820

GCCGGGTCCT TCATCCGGGT CCATAGAGCA GCACGACCAA GCTAACAACG ACAACAATAT 5880

GTGGAATAAT CGCGAGTACT ATGGTAACAG CCCATCGTAT GCAGGCGGAT ACACGCAAGA 5940

TGGCAATATC CACGAGCAGC AACAACACGA TGAGTACACG AGTAATGCGT CATATGGCGG 6000

AAATCAAGGA GCAGGCGGAG GCAGCGGCGG CGGTGGCGGT CTCCGAGTTG CAAATCGTGA 6060

CAGCTCCGAC AGCGAGGGTG CAGATGACGA CGCTTGGAGA CGTGATGCCC TTGCTCAGAT 6120

CAATTTTGCG GGCGGCGCTG CTGCTGCCTC CGCTGGAGCA CCTGCTGCTG GTGCTTCTTC 6180

TTCGCAGCCG GGCCATGCGC AGTAGACGGG ATATGCGTGA GTTTTTTTTT AAATTTCGTA 6240

CATAGAGACC GTTGTATACG CAGGTTTCAA ATTAGAAGAG CGAATATGCA TATCAGCTGT 6300

TGTTCAATGT TCTAGTTTGG GAAGGTTAAC CCCCCCCCCT TCCCCTTCCA AGACTTTTCA 6360

CTTGTTTGTG TGTGATTTAA ATCTGGAGAT TTCAAATCTA CATCTCGCTA TACATAGGTG 6420

TTGTTTGATA ACGTAGGGGG CAGAAGGGTA TCTCGTGATA TTAGACTGGG AGTTGCATGA 6480

ATCAAGGTGT TGAGCAAAAA AAGAGAGAGC GGTGAAGGGC GGGGGGGATA GGTGGTGTGC 6540

ACGTGGCTG 6549

In a preferred embodiment, the plant pathogen is *Cochliobolus heterostrophus*. In another preferred embodiment, the plant pathogen is

Cochliobolus carbonum, C. victoriae, C. sativus, C. specifer, C. homomorphus, C. dactyloctenii, Setosphaeria turcica, S. rostrata, or *Bioplaris sacchari.*

The peptide synthetase of SEQ. ID. No. 2 has a deduced amino acid sequence corresponding to SEQ. ID. No. 3 as follows:

Met Leu Glu Val Asn Gln Gly Tyr Phe Ser Asp Phe Thr Gly Gln 15

Gln Met Gln Asp Asn Arg Asp Ser Tyr Gly Gly Pro Asn Arg Tyr 30

Ser Ser Gly Asp Ala Phe Ser Pro Thr Ala Ala Ile Pro Pro Pro 45

Met Met Asn Pro Asn Asp Leu Pro Leu Gly Ala Ala Glu Thr Met 60

Met Pro Leu Glu Pro Arg Asp Leu Pro Phe Asp Val Tyr Asp Pro 75

His Asn Pro Asn Val Lys Met Ser Lys Phe Asp Asn Ile Gly Ala 90

Val Leu Arg His Arg Ser Arg Thr Gln Pro Arg Thr Thr Ala Phe 105

Trp Val Leu Asp Ala Lys Gly Lys Glu Thr Ala Ser Ile Thr Trp 120

Glu Lys Val Ala Ser Arg Ala Glu Lys Val Ala Lys Val Ile Arg 135

Asp Lys Ser Asn Leu Tyr Arg Gly Asp Arg Val Ala Leu Val Tyr 150

Arg Asp Thr Glu Ile Ile Asp Phe Val Val Ala Leu Met Gly Cys 165

Phe Ile Ala Gly Val Val Ala Val Pro Ile Asn Ser Val Asp Asp 180

Tyr Gln Lys Leu Ile Leu Leu Leu Thr Thr Thr Gln Ala His Leu 195

Ala Leu Thr Thr Asp Asn Asn Leu Lys Ala Phe His Arg Asp Ile 210

Ser Gln Asn Arg Leu Lys Trp Pro Ser Gly Val Glu Trp Trp Lys 225

Thr Asn Glu Phe Gly Ser His His Pro Lys Lys His Asp Asp Thr 240

Pro Ala Leu Gln Val Pro Glu Val Ala Tyr Ile Glu Phe Ser Arg 255

Ala Pro Thr Gly Asp Leu Arg Gly Val Val Leu Ser His Arg Thr 270

Ile Met His Gln Met Ala Cys Ile Ser Ala Met Ile Ser Thr Ile 285

Pro Thr Asn Ala Gln Ser Gln Asp Thr Phe Ser Thr Ser Leu Arg 300

Asp Ala Glu Gly Lys Phe Val Ala Pro Ala Pro Ser Arg Asn Pro 315

Thr Glu Val Ile Leu Thr Tyr Leu Asp Pro Arg Glu Ser Ala Gly 330

Leu Ile Leu Ser Val Leu Phe Ala Val Tyr Gly Gly His Thr Thr 345

Val Trp Leu Glu Thr Ala Thr Met Glu Thr Pro Gly Leu Tyr Ala 360

His Leu Ile Thr Lys Tyr Lys Ser Asn Ile Leu Leu Ala Asp Tyr 375

Pro Gly Leu Lys Arg Ala Ala Tyr Asn Tyr Gln Gln Asp Pro Met 390

Ala Thr Arg Asn Phe Lys Lys Asn Thr Glu Pro Asn Phe Ala Ser 405

Val Lys Ile Cys Leu Ile Asp Thr Leu Thr Val Asp Cys Glu Phe 420

His Glu Ile Leu Gly Asp Arg Tyr Phe Arg Pro Leu Arg Asn Pro 435

Arg Ala Arg Glu Leu Ile Ala Pro Met Leu Cys Leu Pro Glu His 450

Gly Gly Met Ile Ile Ser Val Arg Asp Trp Leu Gly Gly Glu Glu 465

Arg Met Gly Cys Pro Leu Ser Ile Ala Val Glu Glu Ser Asp Asn 480

Asp Glu Asp Asp Thr Glu Asp Lys Tyr Ala Ala Ala Asn Gly Tyr 495

Ser Ser Leu Ile Gly Gly Gly Thr Thr Lys Asn Lys Lys Glu Lys 510

Lys Lys Lys Gly Pro Thr Glu Leu Thr Glu Ile Leu Leu Asp Lys 525

Glu Ala Leu Lys Met Asn Glu Val Ile Val Leu Ala Ile Gly Glu 540

Glu Ala Ser Lys Arg Ala Asn Glu Pro Gly Thr Met Arg Val Gly 555

Ala Phe Gly Tyr Pro Ile Pro Asp Ala Thr Leu Ala Ile Val Asp 570

Pro Glu Thr Ser Leu Leu Cys Ser Pro Tyr Ser Ile Gly Glu Ile 585

Trp Val Asp Ser Pro Ser Leu Ser Gly Gly Phe Trp Gln Leu Gln 600

Lys His Thr Glu Thr Ile Phe His Ala Arg Pro Tyr Arg Phe Val 615

Xaa Gly Ser Pro Thr Pro Gln Leu Leu Glu Leu Glu Phe Leu Arg 630

Thr Gly Leu Leu Gly Phe Val Val Glu Gly Lys Ile Phe Val Leu 645

Gly Leu Tyr Glu Asp Arg Ile Arg Gln Arg Val Glu Trp Val Glu 660

Asn Gly Gln Leu Glu Ala Glu His Arg Tyr Phe Phe Val Gln His 675

Leu Val Thr Ser Ile Met Lys Ala Val Pro Lys Ile Tyr Asp Cys 690

Ser Ser Phe Asp Ser Tyr Val Asn Gly Glu Tyr Leu Pro Ile Ile 705

Leu Ile Glu Thr Gln Ala Ala Ser Thr Ala Pro Thr Asn Pro Gly 720

Gly Pro Pro Gln Gln Leu Asp Ile Pro Phe Leu Asp Ser Leu Ser 735

Glu Arg Cys Met Glu Val Leu Tyr Gln Glu His His Leu Arg Val 750

Tyr Cys Val Met Ile Thr Ala Pro Asn Thr Leu Pro Arg Val Ile 765

Lys Asn Gly Arg Arg Glu Ile Gly Asn Met Leu Cys Arg Arg Glu 780

Phe Asp Asn Gly Ser Leu Pro Cys Val His Val Lys Phe Gly Ile 795

Glu Arg Ser Val Gln Asn Ile Ala Leu Gly Asp Asp Pro Ala Gly 810

Gly Met Trp Ser Phe Glu Ala Ser Met Ala Arg Gln Gln Phe Leu 825

Met Leu Gln Asp Lys Gln Tyr Ser Gly Val Asp His Arg Glu Val 840

Val Ile Asp Asp Arg Thr Ser Thr Pro Leu Asn Gln Phe Ser Asn 855

Ile His Asp Leu Met Gln Trp Arg Val Ser Arg Gln Ala Glu Glu 870

Leu Ala Tyr Cys Thr Val Asp Gly Arg Gly Lys Glu Gly Lys Gly 885

Val Asn Trp Lys Lys Phe Asp Gln Lys Val Ala Gly Val Ala Met 900

Tyr Leu Lys Asn Lys Val Lys Val Gln Ala Gly Asp His Leu Leu 915

Leu Met Tyr Thr His Ser Glu Glu Phe Val Tyr Ala Val His Ala 930

Cys Phe Val Leu Gly Ala Val Cys Ile Pro Met Ala Pro Ile Asp 945

Gln Asn Arg Leu Asn Glu Asp Ala Pro Ala Leu Leu His Ile Leu 960

Ala Asp Phe Lys Val Lys Ala Ile Leu Val Asn Ala Asp Val Asp 975

His Leu Met Lys Ile Lys Gln Val Ser Gln His Ile Lys Gln Ser 990

Ala Ala Ile Leu Lys Ile Ser Val Pro Asn Thr Tyr Ser Thr Thr 1005

Lys Pro Pro Lys Gln Ser Ser Gly Cys Arg Asp Leu Lys Leu Thr 1020

Ile Arg Pro Ala Trp Ile Gln Ala Gly Phe Pro Val Leu Val Trp 1035

Thr Tyr Trp Thr Pro Asp Gln Arg Arg Ile Ala Val Gln Leu Gly 1050

His Ser Gln Ile Met Ala Leu Cys Lys Val Gln Lys Glu Thr Cys 1065

Gln Met Thr Ser Thr Arg Pro Val Leu Gly Cys Val Arg Ser Thr 1080

Ile Gly Leu Gly Phe Leu His Thr Cys Leu Met Gly Ile Phe Leu 1095

Ala Ala Pro Thr Tyr Leu Val Ser Pro Val Asp Phe Ala Gln Asn 1110

Pro Asn Ile Leu Phe Gln Thr Leu Ser Arg Tyr Lys Ile Lys Asp 1125

Ala Tyr Ala Thr Ser Gln Met Leu Asp His Ala Ile Ala Arg Gly 1140

Ala Gly Lys Ser Met Ala Leu His Glu Leu Lys Asn Leu Met Ile 1155

Ala Thr Asp Gly Arg Pro Arg Val Asp Val Tyr Gln Arg Val Arg 1170

Val His Phe Ala Pro Ala Asn Leu Asp Pro Thr Ala Ile Asn Thr 1185

Val Tyr Ser His Val Leu Asn Pro Met Val Ala Ser Arg Ser Tyr 1200

Met Cys Ile Glu Pro Val Glu Leu His Leu Asp Val His Ala Leu 1215

Arg Arg Gly Leu Val Met Pro Val Asp Pro Asp Thr Glu Pro Asn 1230

Ala Leu Leu Val Gln Asp Ser Gly Met Val Pro Val Ser Thr Gln 1245

Ile Ser Ile Val Asn Pro Glu Thr Asn Gln Leu Cys Leu Asn Gly 1260

Glu Tyr Gly Glu Ile Trp Val Gln Ser Glu Ala Asn Ala Tyr Ser 1275

Phe Tyr Met Ser Lys Glu Arg Leu Asp Ala Glu Arg Phe Asn Gly 1290

Arg Thr Ile Asp Gly Asp Pro Asn Val Arg Tyr Val Arg Thr Gly 1305

Asp Leu Gly Phe Leu His Ser Val Thr Arg Pro Ile Gly Pro Asn 1320

Gly Ala Pro Val Asp Met Gln Val Leu Phe Val Leu Gly Ser Ile 1335

Gly Asp Thr Phe Glu Val Asn Gly Leu Asn His Phe Ser Met Asp 1350

Ile Glu Gln Ser Val Glu Arg Cys His Arg Asn Ile Val Pro Gly 1365

Gly Cys Ala Val Phe Gln Ala Gly Gly Leu Val Val Val Val Val 1380

Glu Ile Phe Arg Arg Asn Phe Leu Ala Ser Met Val Pro Val Ile 1395

Val Asn Ala Ile Leu Asn Glu His Gln Leu Val Ile Asp Ile Val 1410

Ser Phe Val Gln Lys Gly Asp Phe His Arg Ser Arg Leu Gly Glu 1425

Lys Gln Arg Gly Lys Ile Leu Ala Gly Trp Val Thr Arg Lys Met 1440

Arg Thr Ile Ala Gln Tyr Ser Ile Arg Asp Pro Asn Gly Gln Asp 1455

Ser Gln Met Ile Thr Glu Glu Pro Gly Pro Arg Ala Ser Met Thr 1470

Gly Ser Met Leu Gly Arg Met Gly Gly Pro Ala Ser Ile Lys Ala 1485

Gly Ser Thr Arg Ala Pro Ser Leu Met Gly Met Thr Ala Thr Met 1500

Asn Asn Leu Ser Leu Thr Gln Gln Gln Gln Gln Gln Tyr Gln Gln 1515

Pro Gly Met Tyr Ala Gln Gln Gln Gly Met His Pro Gln Gln Gln 1530

His Gln Phe Ser Met Ser Asn Thr Pro Pro Gln Gly Pro Pro Gln 1545

Gly Val Glu Leu His Asp Pro Ser Asp Arg Thr Pro Thr Asp Asn 1560

Arg His Ser Phe Leu Ala Asp Pro Arg Met Gln Asn Gln Gly Gln 1575

Met Asn Glu Thr Gly Ala Tyr Glu Pro Met Asn Tyr Gln Asn Ala 1590

Tyr His Pro His Gln Gln Gln Tyr Glu Ser Glu Asp Gly Gly Ser 1605

Arg Leu Ser Gly Pro Val Pro Asp Val Leu Arg Pro Gly Pro Ser 1620

Ser Gly Ser Ile Glu Gln His Asp Gln Ala Asn Asn Asp Asn Asn 1635

Met Trp Asn Asn Arg Glu Tyr Tyr Gly Asn Ser Pro Ser Tyr Ala 1650

Gly Gly Tyr Thr Gln Asp Gly Asn Ile His Glu Gln Gln Gln His 1665

Asp Glu Tyr Thr Ser Asn Ala Ser Tyr Gly Gly Asn Gln Gly Ala 1680

Gly Gly Gly Ser Gly Gly Gly Gly Gly Leu Arg Val Ala Asn Arg 1695

Asp Ser Ser Asp Ser Glu Gly Ala Asp Asp Asp Ala Trp Arg Arg 1710

Asp Ala Leu Ala Gln Ile Asn Phe Ala Gly Gly Ala Ala Ala Ala 1725

Ser Ala Gly Ala Pro Ala Ala Gly Ala Ser Ser Ser Gln Pro Gly 1740

His Ala Gln 1743

In another embodiment of the invention, there is provided an isolated nucleic acid molecule encoding a CPS1 peptide synthetase homolog and which hybridizes to a nucleic acid molecule having a sequence corresponding to SEQ ID No. 41 as follows :

AAGAAGAAAG GGCCGACCGA GTTGACCGAA ATATTGCTAG ATAAGGAAGC ACTGAAGCTG 60

AACGAAGTTG TTGTTTTGGC CATTGGAGAG GAAGTGAGCA AGCGTGTCAA CGAACCCGGC 120

ACTATGAGAG TCGGTGCTTT TGGCTACCCG ATACCAGATG CGACGCTGGC CGTCGTCGAT 180

CCGGAAACTA ATCTTTTGTG TTCACCCTAT TCCATAGGAG AGATCTGGGT AGACTCGCCA 240

TCATTGTCCG GAGGGTTTTG GCAGCTGCAG AAGCACACTG AGACTATTTT CCACGCTCGG 300

CCATATCGTT TCGTAGAGGG CAGCCCAACC CCGCAACTAC TCGAACTGGA GTTTCTACGC 360

ACTGGACTGC TCGGATGCGT GGTAGAAGGC AAAATCTTCG TATTAGGCCT GTACGAGGAC 420

CGGATTAGGC AGCGCGTTGA ATGGGTAGAG CACGGTCAGC TAGAAGCCGA ACATAGGTAT 480

TTCTTCGTGC AGCATCTTGT CACCAGCATT ATGAAAGCTG TTCCAAAGAT TTACGACTGG 540

TAAGTGCTAT CGAATCTCTG GGTAATCAAC CTAACATTGC GCAGCTCGTC TTTCGATTCC 600

TATGTCAACG GCGAATACTT ACCAATCATC CTTATCGAGA CACAGGCCGC ATCAACTGCT 660

CCCACAAATC CAGGCGGGCC ACCACAACAA CTTGACATTC CTTTCCTAGA CTCTCTTTCT 720

GAGCGATGTA TGGAGGTACT GTATCAAGAA CACCACCTTC GGGTGTATTG TGTGATGATC 780

ACTGCACCGA ACACACTCCC GCGAGTCATC AAGAACGGTC GACGAGAAAT TGGAAACATG 840

CTTTGCCGGA GAGAATTTGA CAATGGCTCG CTACCCTGCG TTCACGTCAA GTTTGGCGTC 900

GAGAGGTCGG TCCAGAATAT TGCGCTAGGT GATGACCCTG CTGGCGGCAT GTGGTCTTAC 960

GAGGCGTCGA TGGCACGCCA GCAGTTCCTG ATGCTTCAAG ATAAGCAGTA CTCTGGAGTA 1020

GATCACAGAG AAGTCGTTAT TGACGACAGA ACGTCGACGC CGCTCAACCA GTTCTCCAAC 1080

ATTCATGACC TTATGCAATG GCGCGTACAA CGACAAGCTG AAGAGCTCGC CTACTGCACG 1140

GTAGATGGTC GAGGTAAAGA GGGCAAAGGC GTCAACTGGA AGAAGTTCGA CCAGAAGGTC 1200

GCAGGTGTCG CCATGTACCT GAAGAACAAG GTCAAGGGTC AGACTGGTGA CCACCTGCTC 1260

TTGATGTACA CCCACTCGGA AGACTTTGTC TATGCCGTAC ACGCGTGTTT CGTCCTTGGA 1320

GCTGTGTGTA TACCCATGGC ACCAATCGAC CAGAACAGGC TAAATGAAGA CGCGCCCGCA 1380

CTACTACATA TCATTGCTGA CTTCAAGGTC AAGGCTATCC TCGTCAATGC TGGCGTAGAC 1440

CACCTGATGA AGGTCAAGCA AGTATCGCAG CACATCAAAC AGTCAGCAGT CATTCTCAAG 1500

ATCAACGTAC CGAATACCTA TAACACCACA AAACCACCTA AGCAGTCTAG TGGTTGCCGC 1560

GATCTTAAGC TCACAATACG ACCTGCTTGG ATACAATCTG GTTTCCCTGT TCTAGTATGG 1620

ACATACTGGA CACCTGACCA GAGACGCATA GCTGTGCAAT TAGGTCATAG CCAAATCATG 1680

GCGCTATGCA AAGTTCAGAA AGAAACGTGC CAGATGACGA GCACACGGCC CGTCCTTGGA 1740

TGTGTTCGTA GCACGATCGG TCTTGGCTTC ATACACACCT GTGTTATGGG TATCTTCCTC 1800

GCAGCGCCAA CTTACCTTGT GTCACCTGTC GATTTTGCGC AAAACCCGAA CATCCTCTTC 1860

CAGACCATGT CGAGATACAA GATCAAGGAC GCGTATGCGA CCAGCCAAAT GCTGGACCAC 1920

GCTATTGCAC GAGGTGCTGG CAAGAACATG GCTCTGCACG AGCTCAAGAA CCTCATGATC 1980

GCGACTGACG GTCGGCCGCG CGTAGACGTC TGTAAGTGTT GCGATCCTGT ATAAGCATCT 2040

GAAATCTAAT TCTTGATAGA CCAGCGTGTG CGAGTACACT TCTCGCCAGC AAGTTTGGAC 2100

CGAACGGCAA TCAATACTGT TTACTCACAC GTACTGAATC CTATGGTCGC ATCGCGGTCA 2160

TACATGTGCA TCGAACCCAT AGAACTACAT CTCGATGTCG GTGCCCTTCG AAGAGGTCTC 2220

ATCATGCCTG TCGACCCAGA CACGGAACCT GGTGCTCTCT TAGTCCAGGA CTCGGGTATG 2280

GTACCAGTTA GTACACAAAT TTCAATCGTG AATCCAGAGA CAAACCAGCT TTGCCTAGTC 2340

GGCGAGTATG GCGAAATCTG GGTCCAACC 2370

Preferably, the CPS1 nucleic acid molecule which hybridizes to a nucleic acid molecule having a sequence as set forth in SEQ ID No. 41 is from the plant pathogen *Altenaria solani*. In another preferred embodiment, the CPS1 gene is from a plant pathogen such as *Alternaia alternatherae, A. alternata, A. amaranthi, A. araliae, A. brassicae, A. brassicicola, A. camelliae, A. cassiae, A. cheiranthi, A. cinerariae, A. gossypii, A. helianthi, A. helianthinficiens, A. mali,* or *A. raphani.*

The peptide synthetase product of SEQ ID NO:41 has a deduced amino acid sequence as follows (SEQ ID NO:42):

KKKGPTELTE ILLDKEALKL NEVVVLAIGE EVSKRVNEPG TMRVGAFGYP IPDATLAVVD 60

PETNLLCSPY SIGEIWVDSP SLSGGFWQLQ KHTETIFHAR PYRFVEGSPT PQLLELEFLR 120

TGLLGCVVEG KIFVLGLYED RIRQRVEWVE HGQLEAEHRY FFVQHLVTSI MKAVPKIYDC 180

SSFDSYVNGE YLPIILIETQ AASTAPTNPG GPPQQLDIPF LDSLSERCME VLYQEHHLRV 240

YCVMITAPNT LPRVIKNGRR EIGNMLCRRE FDNGSLPCVH VKFGVERSVQ NIALGDDPAG 300

GMWSYEASMA RQQFLMLQDK QYSGVDHREV VIDDRTSTPL NQFSNIHDLM QWRVQRQAEE 360

LAYCTVDGRG KEGKGVNWKK FDQKVAGVAM YLKNKVKGQT GDHLLLMYTH SEDFVYAVHA 420

CFVLGAVCIP MAPIDQNRLN EDAPALLHII ADFKVKAILV NAGVDHLMKV KQVSQHIKQS 480

AVILKINVPN TYNTTKPPKQ SSGCRDLKLT IRPAWIQSGF PVLVWTYWTP DQRRIAVQLG 540

HSQIMALCKV QKETCQMTST RPVLGCVRST IGLGFIHTCV MGIFLAAPTY LVSPVDFAQN 600

PNILFQTMSR YKIKDAYATS QMLDHAIARG AGKNMALHEL KNLMIATDGR PRVDVYQRVR 660

VHFSPASLDR TAINTVYSHV LNPMVASRSY MCIEPIELHL DVGALRRGLI MPVDPDTEPG 720

ALLVQDSGMV PVSTQISIVN PETNQLCLVG EYGEIWVQ 760

In another embodiment of the invention, there is provided an isolated nucleic acid molecule encoding a CPS1 peptide synthetase homolog and which hybridizes to a nucleic acid molecule having a sequence corresponding to SEQ ID No. 43 as follows :

CTCGAGGTTA GTAAAAGATC CCCGTTTGTT CCACAAATCT CCATCTCCCT CTCAATGCCT 60

TTCTTGGCGC CTCAACCCGC TATTTTGAAG ACAGTTTGTT GTTGTCGCAT GCGACCAAAA 120

ATCATCCTCT CAAGTTTTCA TCGCTGACCT GTTTCTTGGC GTAGGAAGGA GATATCACAC 180

AGAAAGGGTA AGCTGCTTTG CGTCCAGAGT ACTTACAATT GCTTCTCAAT TACTTACGCG 240

CCGGCAGCTA CCAAAAGCGA CGAACTCAAC TTTTCTCCCA ATTCCTCGGT GCACCTCCAC 300

CTCAGATTGC TGCTCTCGCC GAGCCTCAGT CTGGCCTACG CATACACTCG CCCGATGACT 360

CCGACCACCC TTCAGGCGAT GGCCATCGCG CTACCGCCTA TGCCGCTCTC GGTAGCAGCA 420

GCGGTCCAAT CCCAGATTCA CCAGACTCAC CTATGTACCG ACCGCACTCT GGTTATGCTC 480

CTTCAGAATC ACCAAGACCT TCTCCAGCAC AACCTCCACC TTCCCTGCTG CGCCCGGGGG 540

GTTCTCTCGC TGGAGGATCG ACCACTGCTC ACCGCGACTC CCTCTTCTTC TCCCCCTCCC 600

ATCTCGAACC TGAAACCCGG ACAGGTACTA TGATGTCGGG CGACTATGCA TTCAGACCCG 660

AGCAGCAAGG CACATATGGC GAATCCCAGC ATCAACAGCA CCAGTTCCAG CAACAGCAAC 720

AGCCACAGCA GCAACAGCAG TACGATGGGC AGCAGTATGA TGGACGAACT ACAACGCTTC 780

TCGATTCGCA AGGATACTTT TCGGATTTTG CGGGACAGCA GCACTATGAT CAGACTCAAA 840

CCGTTGAGTA TGTGGGACCT CAGCAGCGGT ATTCTTCCAG CGATGCATTC TCTCCAACCG 900

CCGCAATGGC ACCTCCAATG CTTACAACCA ACGACCTCCC ACCGCCGGAA GCGCTTGAGT 960

ACCAGCTGCC CCTTGACCCT CGCGAGGTAC CATTCGCTAT TCAAGATCCC CATGATGATT 1020

CTACGCCAAT GTCAAAGTTC GATAACATCG CAGCTGTACT CAGACATAGA GGCCGAACGA 1080

TTGCTAAGAA GCCGGCATAC TGGGTGTTGG ATAGTAAGGG CAAGGAGATT GCATCGATTA 1140

CGTGGGATAA GCTGGCATCT AGAGCCGAAA AGGTTGCGCA AGTCATCCGC GACAAAAGCT 1200

CTCTGTACCG GGGTGATCGG GTTGCTCTCA TCTACCGCGA TTCAGAGGTT ATTGATTTCG 1260

CCATTGCCTT GCTGGGATGC TTCATTGCTG GAGTTGTTGC CGTTCCCATC AATGATCTGC 1320

AGGACTACCA ACGCTTGAAC CACATTCTTA CTACAACGCA GGCCCATCTA GCGCTGACCA 1380

CCGATAACAA CCTCAAAGCC TTTCAACGAG ACATTACTAC ACAAAAGTTG ACATGGCCAA 1440

AGGGTGTCGA ATGGTGGAAG ACAAACGAGT TTGGCAGTTA TCACCCCAAG AAGAAGGAGG 1500

ATGTCCCGGC TTTGGTTGTT CCCGATCTGG CATATATCGA GTTTTCGCGG GCCCCAACTG 1560

GAGACTTGAG AGGTGTTGTT CTGAGCCACC GAACCATTAT GCACCAAATG GCTTGTCTTA 1620

GTGCGATTAT TTCTACTATC CCGGGTAATG GACCTGGCGA CACTTTCAAC CCGTCTCTTC 1680

GCGACAAGAA TGGTCGACTT ATTGGTGGCG GCGCAAGCAG CGAAATTTTG GTGTCGTACC 1740

TCGATCCCCG TCAGGGCATT GGCATGATTC TGAGCGTGCT ACTGACCGTC TACGGCGGCC 1800

ACACCACTGT TTGGTTCGAC AACAAAGCTG TTGATGTTCC TGGACTGTAC GCCCACCTCC 1860

TTACCAAGTA CAAATCGACC ATCATGATTG CCGACTACCC AGGATTGAAG CGAGCCGCCT 1920

ACAACTACCA GCAAGAGCCA ATGGTGACCC GAAATTTTAA GAAGGGAATG GAGCCAAACT 1980

TTCAAATGAT CAAGCTTTGC TTGATTGACA CCTTGACTGT AGACAGCGGG TCCCACGAAG 2040

TTTTGGCTGA CCGATGGCTA CGACCGTTGA GAAACCCTCG TGCCCGTGAG GTTGTCGCAC 2100

CTATGCTTTG TCTACCTGAA CACGGAGGCA TGGTGATTAG TGTGCGTGAC TGGCTAGGAG 2160

GAGAAGAGCG CATGGGATGC CCATTAAAGC TTGAACTTGG GGAGGATACA GAGTCTGACG 2220

AAGAGAAAGA GGAAACAGAG AAGCCAGCAG TTTCCAATGG CTTTGGTAGT CTCTTGTCAG 2280

GTGGTGGCAC AGCAACAACC GAAGAGAGGG CAAAGAATGA GCTTGGCGAA GTCCTTTTGG 2340

ATCGTGAGGC TCTAAAGACC AACGAAGTTG TGGTGGTGGC CATAGGTAAC GATGCCCGTA 2400

AAAGGGTGAC GGATGACCCA GGCTTGGTAC GGGTCGGTTC TTTTGGATAC CCCATACCCG 2460

ATGCCACACT CTCCGTCGTC GATCCAGAAA CGGGTTTACT GGCGTCACCA CATTCCGTGG 2520

GTGAAATCTG GGTCGACTCC CCTTCTCTTT CAGGTGGTTT CTGGGCGCAG CCAAAGAATA 2580

CTGAGCTGAT TTTCCATGCT CGTCCTTACA AGTTTGACCC AGGTGATCCT ACACCGCAGC 2640

CCGTCGAGCC CGAATTCCTG CGAACAGGCT TGCTGGGCAC CGTCATCGAG GGTAAAATCT 2700

TTGTTCTGGG CCTTTACGAA GACCGAATTC GACAAAAGGT TGAGTGGGTT GAGCATGGAC 2760

ACGAACTAGC AGAGTACCGC TACTTCTTTG TTCAGCACAT CGTTGTGAGC ATTGTCAAGA 2820

ACGTTCCAAA GATATACGAT TGTTCAGCCT TTGACGTCTT TGTCAATGAC GAACACCTGC 2880

CAGTCGTGGT GCTGGAGTCA GCAGCTGCGT CAACGGCACC ATTGACATCT GGAGGACCTC 2940

CTCGACAACC GGATACAGCT CTGCTAGAGT CATTGGCTGA GCGCTGCATG GAGGTTCTCA 3000

TGTCAGAGCA TCATCTGAGA CTGTACTGCG TTATGATCAC AGCACCCGAC ACTTTGCCTC 3060

GAGTTGTTAA GAACGGACGA CGCGAAATTG GTAACATGCT TTGCCGTCGG GAGTTTGATC 3120

TCGGCAACCT TCCATGTGTG CACGTCAAGT TTGGCGTGGA GCATGCAGTA CTTAACCTCC 3180

CTATTGGTGT AGACCCTATA GGTGGTATCT GGTCACCGTT GGCGTCCGAT TCTCGTGCCG 3240

AATTCTTATT GCCAGCTGAC AAGCAATACT CTGGTGTCGA CAGGCGCGAA GTCGTTATCG 3300

ATGACCGTAC TTCAACGCCC CTAAACAATT TCTCTTGCAT TTCGGATCTT ATCCAATGGC 3360

GCGTGGCCCG TCAACCAGAA GAGCTAGCGT ACTGCACAAT CGATGGCAAA AGCCGAGAAG 3420

GTAAGGGTGT AACATGGAAG AAATTCGACA CCAAGGTCGC TTCCGTTGCC ATGTACCTGA 3480

AGAACAAGGT CAAGGTGAGG CCGGGAGACC ACATCATCCT CATGTACACA CATTCAGAGG 3540

AGTTTGTCTT TGCCATCCAT GCCTGCATTT CCTTGGGCGC AATTGTCATT CCCATCGCAC 3600

CCCTCGACCA GAACCGATTG AACGAAGATG TCCCAGCTTT CCTGCATATT GTATCTGATT 3660

ACAACGTCAA GGCTGTGCTG GTCAACGCTG AGGTCGATCA TCTAATCAAG GTAAAGCCTG 3720

TGGCTAGCCA TATCAAACAG TCAGCCCAGG TTCTCAAGAT CACGAGCCCT GCCATCTACA 3780

ACACAACTAA GCCGCCAAAG CAAAGTAGTG GATTGAGGGA TTTGAGATTC ACCATTGACC 3840

CTGCCTGGAT TCGGCCTGGC TACCCCGTCA TTGTTTGGAC TTATTGGACC CCCGATCAAC 3900

GACGAATTTC AGTTCAGCTT GGACATGACA CCATTATGGG CATGTGCAAG GTTCAAAAGG 3960

AAACTTGCCA AATGACAAGT TCAAGACCTG TGCTTGGATG TGTACGAAGC ACGACTGGCC 4020

TAGGCTTTAT TCATACGGCT CTGATGGGAA TTTATATCGG AACACCAACC TACCTCCTAT 4080

CACCTGTCGA GTTTGCAGCC AACCCCATGT CTCTATTCGT CACCTTGTCG AGATACAAGA 4140

TTAAGGATAC TTATGCGACA CCACAGATGC TTGATCATGC CATGAACTCC ATGCAGGCCA 4200

AGGGCTTTAC ACTTCATGAA CTTAAGAACA TGATGATCAC TGCCGAGAGC CGACCAAGAG 4260

TTGATGTTTT CCAAAAGGTC AGACTTCACT TTGCTGGGGC TGGGCTCGAT AGAACTGCTA 4320

TTAACACGGT CTATTCGCAT GTCCTCAACC CCATGGTAGC GTCGCGATCT TATATGTGCA 4380

TCGAGCCTAT TGAGCTTTGG TTGGACACGC AAGCGCTTCG ACGTGGTCTG GTTATTCCTG 4440

TGGACCCTGA ATCAGATCCT CTGGCCCTAC TGGTACAGGA CAGCGGTATG GTTCCAGTTT 4500

CAACCCAAAT AGCCATCATC AACCCTGAAA GCAGAATACA CTGCCTCGAT GGTGAGTATG 4560

GTGAAATTTG GGTCGACTCT GAAGCCTGCG TCAAGTCATT CTATGGCTCC AAAGACGCTT 4620

TTGACGCTGA GCGCTTTGAT GGCCGAGCTC TTGACGGCGA TCCCAACATT CAGTATATCC 4680

GTACCGGAGA CTTGGGTTTC CTTCATAATG TTAGTCGACC TATTGGCCCT AATGGTGCCC 4740

AGGTGGACAT GCAAGTGTTG TTTGTTCTCG GCAACATTGG CGAGACTTTT GAGATCAACG 4800

GATTGAGCCA TTTCCCAATG GATATTGAGA ACTCGGTGGA AAAATGCCAC AGAAACATTG 4860

TGGCGAATGG CTGGTAAGTA TAAAATCTCT ATTTGAAGCG AATATGCTAA CAAAGTCAGT 4920

GCGGTGTTCC AAGCTGGTGG CTTGGTGGTT GTTCTGGTTG AAGTCAACCG CAAGCCATAC 4980

CTGGCATCGA TTGTTCCCGT CATTGTCAAC GCTATCCTCA ATGAACACCA AATCATTGTA 5040

GATATCGTCG CATTCGTCAA CAAGGGAGAC TTCCCACGGT CTCGTCTAGG AGAGAAGCAG 5100

CGTGGCAAGA TTCTTGGTGG CTGGGTTAGT AGAAAGCTGA GGACTCTTGC CCAGTTCTCG 5160

ATTCGCGATA TGGACGCCGA ATCCACAGCT GGTGATATGA TGGATCCTTC TAGAGCATCA 5220

ATGGTCAGCG TACGAAGCGG AGGCGGTGCT GCTCCCGGAT CTTCTAGTTT GAGGAATGTC 5280

GAACCTGCGC CTCAAATCTT GGAGGAGGAA CATGACCAGA TGACTCCTCG TCACGAATAC 5340

```
GAAGCAGCCC CTACCATGAT TTCTGAACTT CCCGACGGCC AAGAGACACC GACAGGGTTT   5400
CAGCACTCGC AATACGAACA CCCACCACAA TCAGCCGGTT CTCAAGCACC AGCCCAGCTG   5460
AACCTTTCTC ACCAGCCCGA TCAAGGATTC GATATGGACT TTTCACGATA TAGTTCAGCA   5520
GAGCCCGATC ACGGCCCTGT CCACAGACGT CCAGTCCCAG GCCAAGCCCA ACAACCCGAG   5580
CCTATGCAAG GGTACGGTCA AGCGCCGCCC CAGATCCGGC TACCAGGTGT TGATGGACGA   5640
GAGGAGGGAG GGTTCTGGTC ACAGCAGGAA AAGAACGAGA AGAGTGAAGA AGACTGGACA   5700
ACTGATGCCA TGATGCATAT GAATCTGGCA GGTGATATGA AACCGCCACG ATGATAATAC   5760
ACAACATAAG AGCGAAGTGA CGAAGCGGAG TCGGAGTTGG GAAGCATTTA GAAACGAATA   5820
ACAAACAATT GGACTTGTCG GTCTGATGGC CTATTTACTT CATTCATAGA TGAGGATTGG   5880
ATAGTGAATA TGTGATTGGA TAAAGCCTGG GTTTGTGAGT TTGTGAATGC AGTGGGTGCT   5940
TGCTATAAGC TGTTTTATTG AGGTCTTTGG AGGAGTGTCT AACAAAGATG CAAAGTTACT   6000
AGT                                                                 6003
```

Preferably, the CPS1 nucleic acid molecule which hybridizes to a nucleic acid molecule having a sequence as set forth in SEQ ID No. 43 is from the plant pathogen *Fusarium graminearium*. In another preferred embodiment, the CPS1 gene is from a plant pathogen such as *Fusarium avenaceum, F. carpineum, F. chlamydosporum, F. coccophilum, F. culmorum, F. episphaeria, F. equiseti, F. flocciferum, F. moniliforme, F. oxysporum, F. redolens, F. sambucinum, F. solani, F. subglutinans, F. trichothecioides, F. udum,* or *F. ventricosum.*

The peptide synthetase product of SEQ ID NO:43 has a deduced amino acid sequence as follows (SEQ ID No. 44):

```
MMSGDYAFRP EQQGTYGESQ HQQHQFQQQQ QPQQQQQYDG QQYDGRTTTL LDSQGYFSDF     60
AGQQHYDQTQ TVEYVGPQQR YSSSDAFSPT AAMAPPMLTT NDLPPPEALE YQLPLDPREV    120
PFAIQDPHDD STPMSKFDNI AAVLRHRGRT IAKKPAYWVL DSKGKEIASI TWDKLASRAE    180
KVAQVIRDKS SLYRGDRVAL IYRDSEVIDF AIALLGCFIA GVVAVPINDL QDYQRLNHIL    240
TTTQAHLALT TDNNLKAFQR DITTQKLTWP KGVEWWKTNE FGSYHPKKKE DVPALVVPDL    300
AYIEFSRAPT GDLRGVVLSH RTIMHQMACL SAIISTIPGN GPGDTFNPSL RDKNGRLIGG    360
GASSEILVSY LDPRQGIGMI LSVLLTVYGG HTTVWFDNKA VDVPGLYAHL LTKYKSTIMI    420
ADYPGLKRAA YNYQQEPMVT RNFKKGMEPN FQMIKLCLID TLTVDSGSHE VLADRWLRPL    480
RNPRAREVVA PMLCLPEHGG MVISVRDWLG GEERMGCPLK LELGEDTESD EEKEETEKPA    540
VSNGFGSLLS GGGTATTEER AKNELGEVLL DREALKTNEV VVVAIGNDAR KRVTDDPGLV    600
RVGSFGYPIP DATLSVVDPE TGLLASPHSV GEIWVDSPSL SGGFWAQPKN TELIFHARPY    660
```

KFDPGDPTPQ PVEPEFLRTG LLGTVIEGKI FVLGLYEDRI RQKVEWVEHG HELAEYRYFF 720

VQHIVVSIVK NVPKIYDCSA FDVFVNDEHL PVVVLESAAA STAPLTSGGP PRQPDTALLE 780

SLAERCMEVL MSEHHLRLYC VMITAPDTLP RVVKNGRREI GNMLCRREFD LGNLPCVHVK 840

FGVEHAVLNL PIGVDPIGGI WSPLASDSRA EFLLPADKQY SGVDRREVVI DDRTSTPLNN 900

FSCISDLIQW RVARQPEELA YCTIDGKSRE GKGVTWKKFD TKVASVAMYL KNKVKVRPGD 960

HIILMYTHSE EFVFAIHACI SLGAIVIPIA PLDQNRLNED VPAFLHIVSD YNVKAVLVNA 1020

EVDHLIKVKP VASHIKQSAQ VLKITSPAIY NTTKPPKQSS GLRDLRFTID PAWIRPGYPV 1080

IVWTYWTPDQ RRISVQLGHD TIMGMCKVQK ETCQMTSSRP VLGCVRSTTG LGFIHTALMG 1140

IYIGTPTYLL SPVEFAANPM SLFVTLSRYK IKDTYATPQM LDHAMNSMQA KGFTLHELKN 1200

MMITAESRPR VDVFQKVRLH FAGAGLDRTA INTVYSHVLN PMVASRSYMC IEPIELWLDT 1260

QALRRGLVIP VDPESDPLAL LVQDSGMVPV STQIAIINPE SRIHCLDGEY GEIWVDSEAC 1320

VKSFYGSKDA FDAERFDGRA LDGDPNIQYI RTGDLGFLHN VSRPIGPNGA QVDMQVLFVL 1380

GNIGETFEIN GLSHFPMDIE NSVEKCHRNI VANGCAVFQA GGLVVVLVEV NRKPYLASIV 1440

PVIVNAILNE HQIIVDIVAF VNKGDFPRSR LGEKQRGKIL GGWVSRKLRT LAQFSIRDMD 1500

AESTAGDMMD PSRASMVSVR SGGGAAPGSS SLRNVEPAPQ ILEEEHDQMT PRHEYEAAPT 1560

MISELPDGQE TPTGFQHSQY EHPPQSAGSQ APAQLNLSHQ PDQGFDMDFS RYSSAEPDHG 1620

PVHRRPVPGQ AQQPEPMQGY GQAPPQIRLP GVDGREEGGF WSQQEKNEKS EEDWTTDAMM 1680

HMNLAGDMKP PR 1692

In another embodiment of the invention, there is provided an isolated nucleic acid molecule encoding a CPS1 peptide synthetase homolog and which hybridizes to a nucleic acid molecule having a sequence corresponding to SEQ ID No. 45 as follows :

AAAAAGAAGG GGCCTACGGA GTTGACCGAG ATATTGCTAG ATAAGGAAGC GCTCAAGATG 60

AACGATGTTG TGGTCCTTGC AATAGGAGAA GAGGCCAGTA AACGTGCGAA TGAGCCTGGC 120

ACAATGCGAG TTGGCGCTTT TGGATACCCA ATACCAGATG CGACGCTAGC CGTCGTAGAT 180

CCAGAGACGA ATCTCTTGTG TTCACCCTAC TCGATAGGAG AGATTTGGGT AGACTCACCT 240

TCATTGTCTG GTGGTTTCTG GCAATTGCAG AAGCACACTG AAACTATATT TCACGCCCGC 300

CCATACCGCT TTGTGGAGGG CAGTCCTACC CCGCAGTTGC TTGAGCTTGA GTTTCTCCGG 360

ACAGGCTTAC TCGGATTCGT CGTAGAGGGC AAGGTCTTTA TCCTTGGTCT CTATGAAGAT 420

CGCATCAGGC AGCGCGTTGA ATGGGTAGAA CATGGTCAGC TGGAAGCTGA ACACAGATAC 480

TTCTTCGTGC AGCACCTCGT CACCAGTATC ATGAAGGCTG TTCCCAAGAT CTACGACTGG 540

TAAGTCTTCT CATGTTTTAG ATGAGCGTTC TAACACTATG CAGCTCATCT TTCGACTCGT 600

ACGTCAATGG CGAATACCTG CCTATCATCC TCATCGAGAC ACAGGCTGCA TCGACAGCCC 660

CTACGAACCC TGGTGGACCG CCACAGCAAC TCGACATCCC CTTCCTAGAC TCACTGTCTG 720

AGCGATGCAT GGAAGTGTTG TATCAAGAAC ACCATCTGCG AGTATACTGC GTCATGATCA 780

CAGCGCCAAA CACATTACCA CGAGTTGTTA AGAATGGTCG ACGAGAAATT GGCAACATGC 840

TCTGTCGAAG AGAATTTGAT AATGGCTCAT TACCTTGTGT CCACGTCAAG TTTGGTGTTG 900

AGAGGTCAGT TCTCAACATC GCGTTGGGTG ATGACCCCTC CGGAGGCATG TGGTCATATG 960

AAGCCTCGAT GGCGCGTCAG CAGTTCTTGA TGCTCCAAGA CAAGCAGTAT TCTGGAGTAG 1020

ATCACCGCGA AGTCGTCATG GATGACAGAA CATCGACACC TCTCAACCAA TTCTCCAACA 1080

TTCACGACCT CATGCAATGG CGCGTATCAC GGCAGGCTGA AGAGCTCGCA TATTGCACAG 1140

TCGACGGTCG AGGCAAAGAA GGCAAGGGCG TCAACTGGAA GAAGTTCGAC CAGAAAGTTG 1200

CGGGTGTCGC AATGTACCTG AAGAACAAGG TCAAAGTGCA AACCGGCGAT CATCTGCTTC 1260

TGATGTATAC GCACTCGGAA GACTTTGTAT ATGCGGTACA TGCATGCTTT GTGCTTGGCG 1320

CTGTATGCAT ACCAATGGCA CCAATCGACC AGAACCGATT GAATGAGGAT GCACCTGCAT 1380

TGCTGCACAT CCTTGCAGAC TTCAAGGTCA AGGCCATCCT CGTCAATGCC GATGTGGATC 1440

ATCTCATGAA GGTCAAGCAA GTATCGCAGC ACATCAAACA ATCAGCAGCC ATCTTCAAGA 1500

TCAACGTGCC GCACACTTAC AACACAACCA AGCCACCTAA GCAGTCGAGT GGTTGTCGGG 1560

ATCTCAAGCT CACAATACGG CCTGCCTGGG TACAGCCTGG TTTCCCAGTT CTTGTATGGA 1620

CATACTGGAC TCCAGATCAA CGCCGTATAG CCGTACAACT AGGTCATAGC CAAATCATGG 1680

CACTAGGCAA GGTCCAGAAG GAGACTTGTC AAATGACAAG TACAAGGCCA GTCCTAGGAT 1740

GTGTACGGAG TACCATCGGA CTTGGCTTCA TTCATACCTG CATCATGGGC ATCTTCCTTG 1800

CCGCACCCAC TTACCTCGTG TCGCCTGTCG ACTTTGCACA AAATCCAAAC ATACTCTTCC 1860

AGACGTTATC AAGATACAAG ATCAAGAATG CGTACGCAAC CAGTCAAATG TTGGATCACG 1920

CTATTGCCCG TGGGGCTGGA AAGAACATGG CCCTGCACGA ACTCAAGAAT CTCATGATTG 1980

CGACTGATGG TAGGCCGCGT GTTGATGTTT ACCAGAGAGT GCGCGTACAC TTTTCACCAG 2040

CAAGCTTGGA CCGGACAGCG ATTAACACAG TCTACTCTCA CGTGCTCAAC CCAATGGTAG 2100

CATCGCGATC ATACATGTGC ATCGAGCCAA TAGAACTGCA TCTCGACGTC AACGCTCTTC 2160

GAAGAGGTCT GATCATGCCC GTCGACCCAG ATACCGAGCC TGGCGCTCTA ATGGTCCAGG 2220

ACTCTGGTAT GGTGCCAGTC TCCACACAAA TAGCAATTGT GAACCCAGAG ACAAACCAGC 2280

TTTGCTTGGT TGGCGAATAT GGCGAAATCT GGGTTCAATC 2320

Preferably, the CPS1 nucleic acid molecule which hybridizes to a nucleic acid molecule having a sequence as set forth in SEQ ID No. 45 is from the plant pathogen *Pyrenophora teres*. In another preferred embodiment, the CPS1

gene is from a plant pathogen such as *Pyrenophora avenae, P. bromi, P. leuceienes, P. phaeocomes, P. schroeteri, P. trichostoma,* or *P. tritici-repentis.*

The peptide synthetase product of SEQ ID No:45 has a deduced amino acid sequence as follows (SEQ ID No. 46):

```
KKKGPTELTE ILLDKEALKM NDVVVLAIGE EASKRANEPG TMRVGAFGYP IPDATLAVVD    60
PETNLLCSPY SIGEIWVDSP SLSGGFWQLQ KHTETIFHAR PYRFVEGSPT PQLLELEFLR   120
TGLLGFVVEG KVFILGLYED RIRQRVEWVE HGQLEAEHRY FFVQHLVTSI MKAVPKIYDC   180
SSFDSYVNGE YLPIILIETQ AASTAPTNPG GPPQQLDIPF LDSLSERCME VLYQEHHLRV   240
YCVMITAPNT LPRVVKNGRR EIGNMLCRRE FDNGSLPCVH VKFGVERSVL NIALGDDPSG   300
GMWSYEASMA RQQFLMLQDK QYSGVDHREV VMDDRTSTPL NQFSNIHDLM QWRVSRQAEE   360
LAYCTVDGRG KEGKGVNWKK FDQKVAGVAM YLKNKVKVQT GDHLLLMYTH SEDFVYAVHA   420
CFVLGAVCIP MAPIDQNRLN EDAPALLHIL ADFKVKAILV NADVDHLMKV KQVSQHIKQS   480
AAIFKINVPH TYNTTKPPKQ SSGCRDLKLT IRPAWVQPGF PVLVWTYWTP DQRRIAVQLG   540
HSQIMALGKV QKETCQMTST RPVLGCVRST IGLGFIHTCI MGIFLAAPTY LVSPVDFAQN   600
PNILFQTLSR YKIKNAYATS QMLDHAIARG AGKNMALHEL KNLMIATDGR PRVDVYQRVR   660
VHFSPASLDR TAINTVYSHV LNPMVASRSY MCIEPIELHL DVNALRRGLI MPVDPDTEPG   720
ALMVQDSGMV PVSTQIAIVN PETNQLCLVG EYGEIWVQ                           758
```

As used herein, the term "nucleic acid" refers to deoxyribonucleic acid (DNA) or ribonucleic acid and polymers thereof in either a single or double stranded form. As used herein, "nucleic acid" also encompasses nucleic acids containing known analogs of naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also encompasses conservatively modified variants thereof such as degenerate codon substitutions and complementary sequences. As used herein, the term "nucleotide sequence" refers to a polymer of DNA or RNA which may be single or double stranded and may contain synthetic, non-natural, or altered nucleotide bases capable of incorporation into DNA or RNA polymers. The terms "nucleic acid" , "nucleic acid molecule", "nucleic acid fragment", or nucleic acid sequence or segment", may also be used interchangeably with the terms "gene", "cDNA", "DNA" and "RNA".

Other DNA molecules of the present invention include DNA molecules that have a nucleic acid sequence which is more than 70% identical to

the nucleotide sequence of SEQ. ID. Nos. 2, 41, 43, or 45. Nucleotide sequence similarity may be determined by the BLAST program with the default parameters (Altschul et al., "Basic Local Alignment Search Tool," J. Mol. Biol., 215:403-410 (1990), which is hereby incorporated by reference).

Preferred sequences include those DNA molecules which will hybridize to a nucleic acid molecule having the sequence of SEQ. ID No. 2, 41, 43, 45 or their compliments. Generally, stringent conditions are selected to be about 50°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. The T_m is dependent upon the solution conditions and the base composition of the probe, and may be calculated using the following equation:

$$T_m = 79.8°C + (18.5 \times \text{Log}[Na+]) + (58.4°C \times \%[G+C]) - (820 / \#bp \text{ in duplex}) - (0.5 \times \% \text{ formamide})$$

More preferred stringent conditions are when the temperature is 20°C below T_m, and the most preferred stringent conditions are when the temperature is 10°C below T_m. Nonspecific binding may also be controlled using any one of a number of known techniques such as, for example, blocking the membrane with protein-containing solutions, addition of heterologous RNA, DNA, and SDS to the hybridization buffer, and treatment with RNase.

Wash conditions are typically performed at or below stringency. Generally, suitable stringent conditions for nucleic acid hybridization assays or gene amplification detection procedures are asas set forth above. More or less stringent conditions may also be selected.

For the purposes of defining the level of stringency, reference can conveniently be made to Sambrook, J., E.F. Fritsch, et al. 1989 "Molecular Cloning: a Laboratory Manual, 2nd Edition, Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press, at 11.45. An example of low stringency conditions is

4-6X SSC/0.1-0.5% w/v SDS at 37°-45° C for 2-3 hours. Depending on the source and concentration of the nucleic acid involved in the hybridization, alternative conditions of stringency may be employed such as medium stringent conditions. Examples of medium stringent conditions include 1-4X SSC/0.25% w/v SDS at $\geq$ 45° C for 2-3 hours. An example of high stringency conditions includes 0.1-1X SSC/0.1% w/v SDS at 60 C for 1-3 hours. The skilled artisan is aware of various parameters which may be altered during hybridization and washing and which will either maintain or change the stringency conditions. For example, another stringent hibridization condition is hybridization at 4X SSC at 65° C, followed by a washing in 0.1X SSC at 65° C for about one hour. Alternatively, an exemplary stringent hybridization condition is in 50% formamide, 4XSSC, at 42° C. Still another example of stringent conditions include hybridization at 62° C in 6X SSC, .05X BLOTTO, and washing at 2X SSC, 0.1% SDS at 62° C.

Other proteins or polypeptides of the present invention include polypeptides that have an amino acid sequence having at least 75 % similarity to the amino acid sequence of at least one of SEQ. ID. No. 3, SEQ ID No. 42, SEQ ID No. 44, or SEQ ID No. 46. In a preferred embodiement of the invention, the protein or polypeptide will have at least 90% similarity with at least one of SEQ. ID No. 3, SEQ ID No. 42, SEQ ID No. 44, or SEQ ID No. 46. Protein sequence similarity may be determined by the BLAST program with the default parameters (Altschul et al., "Basic Local Alignment Search Tool," J. Mol. Biol., 215:403-410 (1990), which is hereby incorporated by reference). The CPS1 protein of SEQ. ID. No. 3 has a molecular weight of about 190-200 kDa, preferably 193.2 kDa. The CPS1 protein contains two structurally similar modules, both of which are similar to SafB1, the first module of saframycin synthetase B (overall 25% identity; 50% similarity) and have apparent amino-acid-activating and thiolation domains with core sequences conserved in known peptide synthetases.

The DNA molecule encoding a CPS1 protein or polypeptide of the present invention can be incorporated in cells using conventional recombinant DNA technology. Generally, this involves inserting the DNA molecule into an expression system to which the DNA molecule is heterologous (i.e., not normally

present). The heterologous DNA molecule is inserted into the expression system or vector in proper sense orientation and correct reading frame. The vector contains the necessary elements for the transcription and translation of the inserted protein-coding sequences. U.S. Patent No. 4,237,224 to Cohen and Boyer, which is hereby incorporated by reference, describes the production of expression systems in the form of recombinant plasmids using restriction enzyme cleavage and ligation with DNA ligase. These recombinant plasmids are then introduced by means of transformation and replicated in unicellular cultures including procaryotic organisms and eukaryotic cells grown in culture. Recombinant genes may also be introduced into viruses, such as vaccinia virus. Recombinant viruses can be generated by transfection of plasmids into cells infected with virus.

Suitable vectors include, but are not limited to, the following viral vectors such as lambda vector system gt11, gt WES.tB, Charon 4, and plasmid vectors such as pBR322, pBR325, pACYC177, pACYC184, pUC8, pUC9, pUC18, pUC19, pLG339, pR290, pKC37, pKC101, SV 40, pBluescript II SK +/- or KS +/- (see "Stratagene Cloning Systems" Catalog (1993) from Stratagene, La Jolla, Calif, which is hereby incorporated by reference), pQE, pIH821, pGEX, pET series (see Studier et. al., "Use of T7 RNA Polymerase to Direct Expression of Cloned Genes," Gene Expression Technology, vol. 185 (1990), which is hereby incorporated by reference), and any derivatives thereof. Suitable vectors are continually being developed and identified. Recombinant molecules can be introduced into cells via transformation, transduction, conjugation, mobilization, or electroporation using methods well known in the art. The DNA sequences are cloned into the vector using standard cloning procedures in the art, as described by Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Springs Laboratory, Cold Springs Harbor, New York (1982), which is hereby incorporated by reference.

A variety of host-vector systems may be utilized to express the protein-encoding sequence(s). Primarily, the vector system must be compatible with the host cell used. Host-vector systems include but are not limited to the following: bacteria transformed with bacteriophage DNA, plasmid DNA, or cosmid DNA; microorganisms such as yeast containing yeast vectors; mammalian

cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus); and plant cells infected by bacteria or transformed via particle bombardment (i.e. biolistics). The expression elements of these vectors vary in their strength and specificities. Depending upon the host-vector system utilized, any one of a number of suitable transcription and translation elements can be used.

Different genetic signals and processing events control many levels of gene expression (e.g., DNA transcription and messenger RNA ("mRNA") translation). Transcription of DNA is dependent upon the presence of a promoter which is a DNA sequence that directs the binding of RNA polymerase and thereby promotes mRNA synthesis. The DNA sequences of eukaryotic promoters differ from those of procaryotic promoters. Furthermore, eukaryotic promoters and accompanying genetic signals may not be recognized in or may not function in a procaryotic system, and, further, procaryotic promoters are not recognized and do not function in eukaryotic cells. Similarly, translation of mRNA in procaryotes depends upon the presence of the proper procaryotic signals which differ from those of eukaryotes. Efficient translation of mRNA in procaryotes requires a ribosome binding site called the Shine-Dalgarno ("SD") sequence on the mRNA. This sequence is a short nucleotide sequence of mRNA that is located before the start codon, usually AUG, which encodes the amino-terminal methionine of the protein. The SD sequences are complementary to the 3'-end of the 16S rRNA (ribosomal RNA) and probably promote binding of mRNA to ribosomes by duplexing with the rRNA to allow correct positioning of the ribosome. For a review on maximizing gene expression, see Roberts and Lauer, Methods in Enzymology, 68:473 (1979), which is hereby incorporated by reference.

Promoters vary in their "strength" (i.e., their ability to promote transcription). For the purposes of expressing a cloned gene, it is desirable to use strong promoters in order to obtain a high level of transcription and, hence, expression of the gene. Depending upon the host cell system utilized, any one of a number of suitable promoters may be used. For instance, when cloning in *E. coli*, its bacteriophages, or plasmids, promoters such as the T7 phage promoter, *lac* promoter, *trp* promoter, *rec*A promoter, ribosomal RNA promoter, the P_R and

P_L promoters of coliphage lambda and others, including but not limited, to *lac*UV5, *omp*F, *bla*, *lpp*, and the like, may be used to direct high levels of transcription of adjacent DNA segments. Additionally, a hybrid *trp-lac*UV5 (*tac*) promoter or other *E. coli* promoters produced by recombinant DNA or other synthetic DNA techniques may be used to provide for transcription of the inserted gene. Bacterial host cell strains and expression vectors may be chosen which inhibit the action of the promoter unless specifically induced. In certain operons, the addition of specific inducers is necessary for efficient transcription of the inserted DNA. For example, the *lac* operon is induced by the addition of lactose or IPTG (isopropylthio-beta-D-galactoside). A variety of other operons, such as *trp*, *pro*, etc., are under different controls.

Specific initiation signals are also required for efficient gene transcription and translation in procaryotic cells. These transcription and translation initiation signals may vary in "strength" as measured by the quantity of gene specific messenger RNA and protein synthesized, respectively. The DNA expression vector, which contains a promoter, may also contain any combination of various "strong" transcription and/or translation initiation signals. For instance, efficient translation in *E. coli* requires a Shine-Dalgarno ("SD") sequence about 7-9 bases 5' to the initiation codon ("ATG") to provide a ribosome binding site. Thus, any SD-ATG combination that can be utilized by host cell ribosomes may be employed. Such combinations include but are not limited to the SD-ATG combination from the *cro* gene or the *N* gene of coliphage lambda, or from the *E. coli* tryptophan E, D, C, B or A genes. Additionally, any SD-ATG combination produced by recombinant DNA or other techniques involving incorporation of synthetic nucleotides may be used.

The present invention also relates to anti-sense nucleic acid for essential cell proteins, such as replication proteins, which serve to render the host cells incapable of further cell growth and division. Anti-sense regulation has been described by Rosenberg et al., "Production of Phenocopies by Kruppel Antisense RNA Injection Into Drosophila Embryos," Nature, 313:703-706 (1985); Preiss et al., "Molecular Genetics of Kruppel, A Gene Required for Segmentation of the Drosophila Embryo," Nature, 313:27-32 (1985); Melton, "Injected Anti-sense

From: Elijah Cocks
To: cohn@psfc.mit.edu
Date: 5/29/02 3:37PM
Subject: Enlarged Volume and Decoupled Air Flow Control

Dan,

I received your voicemail message about changing the MIT case numbers for Enlarged Volume and Decoupled Air Flow Control. These cases were originally input into our system as being Continuation in Parts (CIPs) of MIT 8530 Low Power Compact Plasma Fuel Converter (hence the designations CIP2 and CIP3). It has become clear that we are no longer claiming priority to this original application with these two applications; however, we should relabel internally with whatever actual MIT Case No. you believe these cases should actually be tagged under. Do you happen to know which MIT Case No. that is?

Also, and perhaps more importantly with respect to filing the Enlarged Volume Plasmatron application, the two DOE grant numbers currently listed in the application were the ones for the MIT 8530 matter. Do you know if there are, in fact, different DOE grant numbers for whatever MIT Case No. the Enlarged Volume Plasmatron should actually be under? Do you happen to know what these DOE Grant numbers might be?

Regards,
Lij

****

Elijah Cocks
Patent Attorney
Choate, Hall & Stewart
Exchange Place
53 State Street
Boston, MA 02109

(617) 248-4792 - Direct Dial
(617) 248-5000 - Main
(617) 248-4000 - Fax

E-mail: ecocks@choate.com

CC: Pasternack, Sam

RNAs Specifically Block Messenger RNA Translation In vivo," Proc. Natl. Acad. Sci. USA, 82:144-148 (1985); Izant et al., "Constitutive and Conditional Suppression of Exogenous and Endogenous Genes by Anti-sense RNA," Science, 229:345-352 (1985); Kim et al., "Stable Reduction of Thymidine Kinase Activity in Cells Expressing High Levels of Anti-sense RNA," Cell, 42:129-138 (1985); Pestka et al., "Anti-mRNA: Specific Inhibition of Translation of Single mRNA Molecules," Proc. Natl. Acad. Sci. USA, 81:7525-7528 (1984); Coleman et al., "The Use of RNAs Complementary to Specific mRNAs to Regulate the Expression of Individual Bacterial Genes," Cell, 37:429-436 (1984); and McGarry et al., "Inhibition of Heat Shock Protein Synthesis by Heat-Inducible Antisense RNA," Proc. Natl. Acad. Sci. USA, 83:399-403 (1986), which are hereby incorporated by reference.

Once the isolated DNA molecules encoding the CPS1 protein or polypeptide, as described above, has been cloned into an expression system, they are ready to be incorporated into a host cell. Such incorporation can be carried out by the various forms of transformation noted above, depending upon the vector/host cell system. Suitable host cells include, but are not limited to, bacteria, virus, yeast, mammalian cells, insect, plant, and the like. In the present invention, the host cells may be from plants such as corn, oat, grass, weed, bamboo, and sugarcane.

One aspect of the present invention involves using an inhibitor of the CPS1 protein to interfere with the plant infection process in order to impart disease resistance to plants. In one mechanism, the peptide synthetase CPS1 could be inhibited by an appropriate drug, thereby causing the plants to be resistant to fungal attack. Alternatively, the CPS1 nonribosomal product could be degraded by an enzyme for which it is a suitable substrate, and, when the gene encoding this enzyme is genetically engineered into plants, the plants will become resistant to fungal attack.

In this aspect of the present invention, large numbers of compounds can be screened for their activity as inhibitors of CPS1 protein by a high-throughput screening assay as described in U.S. Patent No. 5,876,946 to

Burbaum et al., which is hereby incorporated by reference. Generally, a library of compounds is assayed for inhibition of an enzyme catalyzed reaction and the amounts of fluorescence bound to individual suspendable solid supports measured to determine the degree of inhibition. For example, the amount of fluorescence bound to a microbead in the presence of inhibitory compounds is greater than for non-inhibitory compounds. The amounts of fluorescence bound to individual beads are determined by confocal microscopy. Using this type of assay, inhibition can be determined of a peptide synthetase such as CPS1. For CPS1, the substrate can be amino acids (or hydroxy acids), linked at one end to the microbead and at the other end to a fluorescent label. The enzyme inhibitors can be utilized to impart fungal resistance to a variety of plants including oats, grasses, weeds, sugarcane, and corn in particular.

Thus, the present invention provides a method for identifying inhibitors of a CPS1 protein , wherein said CPS1 protein is a peptide synthetase of a plant pathogen. The method comprises: providing a CPS1 protein or polypeptide, contacting the protein or polypeptide with potential inhibitor compounds; determining peptide synthetase activity, and selecting compounds which decrease the peptide synthetase activity. The method is especially useful in identifying inhibitors of a CPS1 protein from plant pathogens of the genera *Cochliobolus, Alternaia, Fusarium*, and *Pyrenophora* such as those described hereinabove. Preferably, the method may be used to identify inhibitors of a CPS1 protein from *Alternaria solani, Fusarium graminearium, and Pyrenophora teres.*

Another aspect of the present invention involves using one or more of the above DNA molecules encoding a CPS1 protein or polypeptide or a gene encoding an enzyme that degrades the CPS1 N.R.P. product to transform plants in order to impart fungal resistance to the plants. This concept of pathogen-derived resistance, according to U.S. Patent No. 5,840,481 to Johnston and Sanford, which is hereby incorporated by reference, is that host resistance to a particular parasite can effectively be engineered by introducing a gene, gene fragment, or modified gene or gene fragment of the pathogen into the host. This approach is based on the fact that in any parasite-host interaction, there are certain parasite-encoded cellular functions (activities) that are essential to the parasite but not to the host

and that when one of the essential functions of the parasite such as survival or reproduction is disrupted, the parasitic process will be stopped. "Disruption" refers to any change that diminishes the survival, reproduction, or infectivity of the parasite. Such essential functions, which are under the control of the parasite's genes, can be disrupted by the presence of a corresponding gene product in the host which is (1) dysfunctional, (2) in excess, or (3) appears in the wrong context or at the wrong developmental stage in the parasite's life cycle. If such faulty signals are designed specifically for parasitic cell functions, they will have little effect on the host. Therefore, the procedure for making plants, for example, resistant to infection by one or more fungus involves isolating DNA coding for a gene such as CPS1 of a fungus, operably linking the DNA within an expression vector, transforming the plant cell or plant tissue with the expression vector, and growing the transformed plant cells or plant tissue in the presence of the fungus such as e.g., *Cochliobolus heterostrophus, Alternaria solani, Fusarium graminearium, or Pyrenophora teres,* where the CPS1 DNA is expressed as a gene product and the CPS protein disrupts the essential activity of the fungi.

Thus, the present invention provides a method of imparting disease resistance to a plant by over-expressing a CPS1 polypeptide in the plant, wherein the polypeptide has protein synthetase activity. In praticing this aspect of the invention, the plant may be any plant in which it is desired to impart disease resistance. Thus, the plant may be an agrigultural crop or ornamental plant. The plant may be herbacious or woody. The plant may be a monocot or dicot. Examples of plants which may be used in practicing the present invention, include but are not limited to, corn, oats, grasses, weeds, sugarcane. barley, wheat, rice, tomato, potato, citrus, malus, rye, cotton, brassica, cabbage, and carrot. Many other plants may also be used in the practice of the present invention. As a guideline, plants which serve as hosts for *Cochliobolus sp., Fusarium sp., Alternaria sp.,* and *Pyrenophera sp.*, may be used. Reference to host plants may be conveniently made to *Fungi on Plants and Plant Products in the United States*, David, F. Farr et al. editors, American Phytopathological Society Press, St. Paul, Minnesota, 1989.

In another preferred embodiment, the CPS1 peptide is from the genera *Cochliobolus, Alternaria, Fusarium* ,or *Pyrenophora.* In an even more preferred embodiment, the CPS1 gene is from *Cochliobolus heterostrophus*, *Alternaria solani*, *Fusarium graminearium* or *Pyrenophora teres.*

Promoters and other regulatory regions which function in plants are well known and include e.g., constitutive promoters, inducible promoters, temporally regulated and tissue specific promoters. Examples of constitutive promters include e.g., actin, CAMV 35S, MAS, ubiquitin, rice cyclophilin, maize H3 histone, and actin 2. Examples of tissue specific promoters include e.g., leaf specific promoters such as the RuBisCo ssu, Cab (chlorophyll a/b/binding) protein, and the AldP gene promoter from rice (Kagaya et al., 1995 *Molecular and General. Genetics 248*:668-674. Examples of root specific promoters include e.g., beta tubulin (Oppenheimer et al. *Gene 63*:87, 1988), and SbPRP1 (Suzuki et al., *Plant Mol. Biol. 21*:109-119, 1993. Thus, using well known methods and widely available regulatory sequences, the skilled artisan is direct expression of a subject CPS nucleic acid molecule in a plant.

To provide regulated expression of a CPS gene of the present invention, plants are transformed with a vector which replicates in a plant cell and which have a promoter which directs expression of the CPS gene product in the plant. Methods of plant transformation are well known in the art. A vector comprising a subject nucleic acid molecule coding for a CPS gene or fragment thereof may be introduced into a plant by leaf disk transforamtion-regeneration procedure as described by Horsh et al. (1985) *Science 227*:1229-1231. Other methods of transformation such as protoplast culture (Horsh et al. 1984 *Science, 223*:496; DeBlock et al. (1984) *Embo J. 2*:2143; Barton et al. (1983) *Cell, 32*:1033) may also be used and are within the scope of this invention.

In transforming dicot plant species, plants may be transformed *Agrobacterium*-derived vectors such as those described in Klett et al. (1987) *Annu. Rev. Plant Physiol., 38*:467. Other well known methods are available to insert the subject CPS genes into plant cells. Such alternative methods include biolistic approaches (Klein et al. 1987, *Nature*:327:70), electroporation,

microinjection (Potrykus and Spangenberg eds., *Gene Transfer to Plants*, Springer Verlag, Berlin , 1995), chemically-induced DNA uptake, the use of viruses or pollen as vectors, liposome mediated transformation, transformation using wounded or enzyme-degraded immature embryos, or wounded or enzyme degraged embryonic callus.

For transformation of monocot plant species, a number of well known methods may be used such as biolistic methods (Tang, K. W. et al. 2000, "Acta Biotechnologica. 20(2): 175-183; de-Villiers, S. M. et al., 2000, South African Journal of Plant and Soil. 17(1): 50- 53) protoplast transformation, electroporation of partially permeabilized cells, and introduction of DNA using glass fibers, etc. *Agrobacterium* may also be used (Raineri, D. M., et al., 1990 Bio/Technology Vol. 8 :33-38).

As used herein, "transformation" refers to the transfer of an exogenous nucleic acid molecule into a host cell. The nucleic acid molecule may be stably or transiently introduced into the host cell and may be maintained non-integrated for example, as a plasmid, or alternatively, may be integrated into the host genome. The resulting transformed plant cell(s) may then be used to regenerate a transformed plant via standard methods.

The invention is further illustrated by the following examples which are not intended in any way to limit the scope of the invention.

EXAMPLES

EXAMPLE 1 -- ***Materials and Methods for Examples 2-7***

Strains, Media, Crosses and Transformation. C4 (*Tox1*$^+$; *MAT-2*) and C5 (*Tox1*$^-$; *MAT-1*) are member of near-isogenic *C. heterostrophus* strains (Leach et al., 1982, which is hereby incorporated by reference). R.C4.2696 (*Tox*$^+$; *MAT-2*; *hygB*^{R}) is a C4-derived mutant generated using the REMI mutagenesis procedure (Lu et al., "Tagged Mutations at the *Tox1* Locus of *Cochliobolus heterostrophus* Using Restriction Enzyme-Mediated Integration," Proc. Natl. Acad. Sci. USA, 91:12649-12653 (1994), which is hereby incorporated by reference). Strains 1301R33 (*Tox*$^-$; *MAT-2*; *hygB*^{R}), 1301R45 (*Tox*$^-$; *MAT-1*;

hygB^R), 1301R26 (*Tox*$^+$; *MAT-2*; *hygB*^{R}) are progeny of the cross C5 X R.C4.2696. Culture media, including CM (complete medium), CMX (complete medium with xylose instead of glucose), CMNS (CM with salts omitted), and MM (minimal medium) have been described, as have mating procedures (Leach et al., 1982; Turgeon et al., "Transformation of the Fungal Maize Pathogen *Cochliobolus heterostrophus* Using the *Aspergillus nidulans amdS* Gene," Mol. Gen. Genet., 201:450-453 (1985), which are hereby incorporated by reference). All strains were grown at 24°C under the warm white light or black light (F40/350BL) (Sylvania Inc., Danvers, MA). Ascospore germination was done at 32°C in the dark for 3 days. REMI transformants were purified by transferring the transformants from the original REMI plates to fresh CMNS medium containing hygromycin B (Calbiochem^R) at 80 ug/ml. For conidiation, stable transformants were transferred to CMX containing the same drug but at a higher concentration (120 ug/ml) to compensate for reduced drug activity due to the inhibition by the salts in the medium. Single conidia were picked up under a dissecting microscope and grown on CMNS hygromycin B plates; stable colonies were then transferred to individual CMX/ hygromycin B plates. All purified transformants were stored at -70°C in CM liquid medium containing 25% of glycerol in 96-well microtiter dishes.

Bioassays. Fungal strains were grown on CMX plates (100 X 15mm) for 7-10 days at 24°C under the light for maximum conidiation. To verifiy normal T-toxin production by a race T isolate, 1.0 ml of T-toxin-sensitive *E. coli* (DH5a) cells were evenly spreaded on LB medium containing ampicillin (100 ug/ml) and the plates were allowed to air dry for 30 min in a laminar hood. Agar plugs bearing fungal mycelia were inoculated (upside down) onto the *E. coli* cell lawn and the plates were incubated at 32°C. Wild type race T and race O were used as controls for each assay plate. T-toxin-producing strains of the fungus will inhibit growth of the *E. coli* cells and produce halos. *Tox*$^-$ mutants can be distinguished from wild type by failure to produce a halo (tight) or by production of halos smaller (leaky) or larger than wild type (overproducing). All *Tox*$^-$ mutants were transferred to Fries medium (Pringle et al., "The Isolation of the Toxin of *Helminthosporium victoriae*," Phytopathology, 47:369-371 (1957),

which is hereby incorporated by reference), which optimizes toxin production, and retested. T-cytoplasm corn plants (inbred W64A) are used to verify the *Tox*$^-$ mutants identified from the *E. coli* assay using the procedure described below. Mutants defective in T-toxin production fail to produce typical race T symptoms on T-corn. Pathogenicity phenotype on N-cytoplasm corn and virulence of *Tox*$^+$ strains to T-cytoplasm corn were determined by a plant assay where about 3,000 transformants generated using the REMI mutagenesis procedure (Lu et al., "Tagged Mutations at the *Tox1* Locus of *Cochliobolus heterostrophus* Using Restriction Enzyme-Mediated Integration," Proc. Natl. Acad. Sci. USA, 91:12649-12653 (1994), which is hereby incorporated by reference) were screened for mutants defective in ability to cause disease on corn plants. Two week old N-cytoplasm corn plants (inbred W64A) grown in the green house (5-6 plants in one 4" X 6" pot) were inoculated with 5 ml conidial suspensions (10^5 conidia/ml) using a pressurized Preval Spray Gun Power Unit thin layer chromatography sprayer (Alltech Associates, Deerfield, IL), incubated in the mist chamber for 24 hours (23°C) and then taken to the growth chamber (23°C, 80% humidity, 14 hours of light). The mutant phenotypes were determined by occurrence of apparent variations in disease symptom development, mainly by lesion size comparison. Mutants producing lesions smaller than wild type were retested and lengths of typical lesions from each mutant were compared with wild type 7 days after inoculation and measurements were taken for statistical evaluation.

DNA manipulations and sequencing. Genomic and plasmid DNA preparation, restriction enzyme digestions, gel electrophoresis and gel blot analysis were done using standard protocols (Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, New York:Cold Spring Harbor Laboratory Press (1989), which is hereby incorporated by reference). DNA was sequenced at the Cornell DNA Sequencing Facility using TaqCycle automated sequencing with DyeDeoxy terminators (Applied Biosystems, Foster City, CA). pUCATPH was used for subcloning (Table 1). Primers used for sequencing (Table 2) were designed using Primer Select (DNASTAR Inc., LaserGene system) and synthesized by the Cornell

Oligonucleotide Synthesis Facility. Sequencing of each plasmid clone was initiated with vector-specific primers or primers designed to previously determined sequences. Sequences obtained were analyzed using the same system and nucleotide or protein database searches were performed with the BLAST program (Altschul et al., "Basic Local Alignment Search Tool," J. Mol. Biol., 215:403-419 (1990), which is hereby incorporated by reference).

Table 1. Transformation vectors and clones used

Plasmid	Length (kb)[a]	Characteristics	Reference
pUCATPH	5.1	See Figure 29.	
pUCATPH N	4.6	Cloning vector, same as pUCATPH but lacking a 420 bp *Nar*I fragment containing the *Hin*dIII site	Figure 8
p214B7	9.2	A clone containing pUCATPH recovered from the tagged site in mutant R.C4.2696 by religation of *Bgl*II-digested genomic DNA	Figure 2
p214M1	6.3	As above but with *Msc*I-digested genomic DNA	Figure 2
p214S1	9.3	As above but with *Sac*I-digested genomic DNA	Figure 2
p214S1N	3.3	*Nar*I fragment derived from 214S1 containing a 0.8 kb *Nar*I-*Sac*I fragment of genomic DNA ligated to pUC18	Figure 4
p214SNP	8.4	Vector for targeted integration constructed by ligating *Hin*dIII-digested pUCATPH into the *Hin*dIII site of p214S1N	Figure 4
p118BSP	7.3	Vector for targeted integration constructed by ligation of a 2.2 kb *Sac*I fragment of p118B14 into the *Sac*I site of pUCATPH	Figure 6
p118BCS	5.4	Vector for targeted integration constructed by ligation of a 0.8 kb *Ssp*I fragment of p118BC4 into the *Ssp*I site of pUCATPHN	Figure 9
p118B14	10.4	A clone recovered from the p214SNP integration site in transformant #118 by ligation of a *Bgl*II-digested genomic DNA fragment containing the entire vector	Figure 5
p118BC4	6.7	A clone recovered from same site as above but by ligation of a *Bcl*I-digested genomic DNA fragment containing part of vector (214SNP) sequence	Figure 5
p9P2	7.3	A clone recovered from the p118BSP integration site in transformant #9 by ligation of a *Pst*I-digested genomic DNA fragment containing pUC18	Figure 7
p12H6	8.0	A clone recovered from the p118BCS integration site in transformant #12 by ligation of a *Hin*dIII-digested genomic DNA fragment containing the entire vector.	Figure 9

a. An underlined kb number indicates that the plasmid carries genomic DNA sequences (see related figures for details).

Table 2. Primers used for sequencing recovered genomic DNA flanking the REMI insertion site at the R.C4 2696 mutation.

Name[a]	Position[b]	Sequence[c]	Plasmid[d]	Origin[e]
M13RMT		GCGGATAACAATTTCACACAGGA SEQ. ID. No. 4	A	pUC18
1. RP1b	775	AGGCCCAGCTGCTTCTCTTG SEQ. ID. No. 5	A	214B7TrpC
2. RP2	604	ACTCGGACCGGACGGAATAACAA SEQ. ID. No. 6	A	214B7RP1b
3. RP3	119	CGGAAGGAGTGCGAACAA SEQ. ID. No. 7	A	214B7RP2
4. RP4	-232	GCTGCTTGCATCTGGTCTTG SEQ. ID. No. 8	A	214B7RP3
5. RP5	-812	AGACCCAGCTGTTGCCCATTG SEQ. ID. No. 9	A	214B7RP4
6. RP5b	-1215	CGGAGACGCAAAGCCTGAGA SEQ. ID. No. 10	A	214B7RP4
7. RP6	-1392	TGCCAGCTGCGTCCAAGAAG SEQ. ID. No. 11	A	214B7RP5
8. RP7	-1839	GCTAGCATGGCCCTCACAC SEQ. ID. No. 12	A	214B7RP6
TrpC		TGTGTTGACCTCCACTAGCTC SEQ. ID. No. 13	A	PUCATPH
9. FP1	1885	CTACGGGATGCAGAGGGAAAGT SEQ. ID. No. 14	A	214B7TrpC
10. FP1b	1828	GCCATGATTAGCACGATACCC SEQ. ID. No. 15	B	214B7TrpC
11. FP2	2028	CGCGCTGCATACAACTACCAA SEQ. ID. No. 16	B	214M1FP1b
12. FP3	2490	TGGTGGCACTACAAAGAACA SEQ. ID. No. 17	C	214M1FP2
13. FP4	2949	CAGCGTGTTGAATGGGTAGAA SEQ. ID. No. 18	C	214S1FP3
14. FP4B	2745	CTGGGTAGATTCGCCTTCAC SEQ. ID. No. 19	C	214S1FP4
15. FP5	3421	GAGCGATCAGTGCAGAACATT SEQ. ID. No. 20	C	214S1FP4
16. FP6	3948	CGCTGACGTTTGACCATCTGA SEQ. ID. No. 21	C	214S1FP5
17. FP7	4411	GCATATGCAACGAGTCAAA SEQ. ID. No. 22	C, D	214S1FP6
18. FP8	5035	ACGGTGCACCTGTTGATA SEQ. ID. No. 23	D	118B14FP7
19. FP9	5457	ATGCGCACAATAGCCCAGTA SEQ. ID. No. 24		118BC4FP8
20. RP48	2865	TTCAAGCAACTGTGGCGTAGG SEQ. ID. No. 25	D	214S1FP6
21. FP10	5790	GATCCTAGCGACCGCACACCAAC SEQ. ID. No. 26	F	9P2FP9
22. FP11	6327	CCTGCTGCTGGTGCTTCT SEQ. ID. No. 27	F	9P2FP10

23. FP11b	6211	GAGTTGCAAATCGTGACAGC SEQ. ID. No. 28	F	9P2FP10
24. FP12	6457	TATCAGCTGTTGTTCAATGTTCTA SEQ. ID. No. 29	F	9P2FP11
25. FP13	6854	TGTTATCCCATTGCCATTG SEQ. ID. No. 30	F	9P2FP12
26. FP14	7400	AAGGACGGAGATTGGTGGAG SEQ. ID. No. 31	F	9P2FP13
27. FP15	7771	GGAGATGGCGGTGACGA SEQ. ID. No. 32	F	9P2FP14
28. FP16	8145	GCATGGCTTGTGGAGGAC SEQ. ID. No. 33	F	9P2FP15
29. FP17	8492	AGATTGTGGCTAGTATGGAGGTAA SEQ. ID. No. 34	F	9P2FP16
M13F40		GTTTTCCCAGTCACGAC SEQ. ID. No. 35	G	pUC18
30. RP1	8953	TACTACTAGCATACCAGCATACCT SEQ. ID. No. 36	G	9P5M13F4
31. RP2	8559	TCAACCTCGGAATACCAAGTC SEQ. ID. No. 37	G	9P5RP1

a. "RP" indicates reverse primer; "FP" indicates forward primer. Primers designed to genomic DNA sequences are numbered in order. For stock tube and the notebook, Primers 1-17 have a leading number "214"; 18-20 with "118"; 21-29 with "9P2" and 30-31 with "9P5". M13RMT(a M13R mutant version; there is a mutation in the polylinker of pUC18) and M13F-40 are provided by Cornell DNA Sequenceing Facility. TrpC primer site is in the pUCATPH *TrpC* promoter region 38 bp from *Sal*I site with sequencing direction from *Sal*I to *Kpn*I.

b. The position of the first base of each primer corresponds to the assembled sequence (*CPS1* + *TES1*, total 11.3 kb found in Lu's folder in the computer Yoder lab1)

c. Each primer sequence is given in the 5' to 3' direction

d. Plasmids used as templates for each sequencing reaction. A = p214B7; B = p214M1; C = p214S1; D = p118B14; E = p118BC4; F = p9P2. G = p9P5 (=9P2)

e. Original sequences that were used for primer design can be found in the CPS1 sequence notebook or in Shunwen Lu's folder (CPS1 sequence) in the computer Yoder lab1 under the same names as listed.

Recovery of tagged DNA from the REMI insertion site and targeted gene disruption. Genomic DNA of mutant R.C4.2696 was digested with *Bgl*II, *Msc*I (no sites in pUCATPH) or *Sac*I (which cuts the vector once) and purified by phenol extraction and ethanol precipitation, then dissolved in TE (pH 8.0). Ligation was performed in 50 ul reaction mixture, containing 1 x T4 DNA ligase buffer with 10 mM ATP, 60 units T4 DNA ligase (New England Biolabs, Beverly, MA) and 3 ug of *Bgl*II-digested genomic DNA, at 14°C overnight. 10 ul of ligation mixture was used to transform 200 ul of competent DH5a cells,

prepared using the calcium chloride treatment (Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, New York:Cold Spring Harbor Laboratory Press (1989), which is hereby incorporated by reference), to ampicillin resistance. Ampicillin resistant clones were analyzed by digestion of plasmid DNA with several diagnostic restriction enzymes and clones containing the REMI vector plus flanking genomic DNA were sequenced using the vector-specific primers (M13R or *TrpC*). Three plasmids, p214B7, p214M1 and p214S1 (Figure 2) were recovered and used for sequencing. For targeted gene disruption in wild type, p214B7 was amplified and plasmid DNA purified by equilibrium centrifugation in CsCl-ethidium bromide gradients (Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, New York:Cold Spring Harbor Laboratory Press (1989), which is hereby incorporated by reference). 30 ug of plasmid DNA (linearized with *Bgl*II for double crossover integration as illustrated in Figure 3) were used to transform wild type and the transformants were purified by isolation of single conidia, assayed for pathogenicity and characterized by gel blot analysis.

Sequence extension by targeted integration and plasmid rescue. Two overlapping cosmid clones were isolated by probing a genomic DNA library of C4 constructed on a cosmid vector, but both extended into the left region only of p214B7. To extend to the right, a chromosome walking strategy was employed (Figure 3). Three targeted gene disruption experiments (each followed by plasmid rescue) were done successively. In the first experiment, a vector was constructed as follows (see Figure 4 for details): p214S1 was digested with *Nar*I and religated to create p214S1N, which was then digested with *Hin*dIII and ligated into the *Hin*dIII site of pUCATPH to create p214SNP for transformation of race O (C5). One transformant (Tx118) resulting from homologous integration (confirmed by gel blot analysis) was used for plasmid rescue as described above. Two new plasmids p118B14 and p118BC4 were recovered, both of which carry sequence at the 3' end but only 172 and 680 bp more than p214S1, respectively (Figure 5). To continue the walk, p118B14 was digested with *Sac*I and ligated into the *Sac*I site of pUCATPH to create p118BSP (Figure 6). This vector was transformed into wild type and one plasmid, p9P2 was recovered (from transformant Tx9), which extends 4.4 kb into the region 3' of p118BC4 and contains the 3' end of *CPS1*

(Figure 7). A third experiment was done in an attempt to recover a 15 kb *Xho*I fragment at the 3' end of that tagged gene. p118BCS (Figure 8) was constructed by subcloning a 0.8 kb *Ssp*I fragment into the same site of pUCATPHN. Plasmid rescue using *Xho*I digested-genomic DNA of a transformant (TX12) failed to recover the 15 kb *Xho*I fragment, but p12H6 was recovered using *Hin*dIII-digested genomic DNA of the same transformant; the genomic DNA matched that already cloned on p9P2 (Figure 9).

EXAMPLE 2 -- ***Characterization of the REMI mutant.***

1) In all culture conditions used in the lab, mutant R.C4.2696 grows just like wild type with no variations in growth rate, color and morphological features (Figure 10A). It produces normal conidia that germinate and form infection structures like wild type when induced on artificial surfaces (Figure 10B) and shows normal mating ability when crossed to wild type testers. No pleiotropic phenotypes associated with the mutation have been detected so far.

2) The mutant differs from wild type in the ability to cause disease on corn plants. When tested on T-cytoplasm corn, the mutant produces race T type symptoms but the disease develops more slowly than with wild type although it produces wild type levels of T-toxin as detected in a microbial assay (Figure 11), suggesting that the reduced virulence is not related to a deficiency in the ability to produce T-toxin. This is clearer on N-cytoplasm corn where the mutant produces lesions significantly smaller than those produced by wild type (Figure 12). When the mutant was crossed to a wild type race O tester, the small lesion phenotype and ability to produce T-toxin segregated independently, indicating that mutant phenotype is not associated with the reduced fitness trait tightly linked with the *Tox1* locus (Klittich et al., "Reduced Fitness Associated With *Tox1* of *Cochliobolus heterostrophus*," Phytopathology, 76:1294-1298 (1986), which is hereby incorporated by reference). The statistical evaluation of lesion size in the race O genetic background indicates that the mutation causes 60% reduction in the fungal virulence to corn plants (Figure 13).

3) The mutant phenotype is caused by a tagged, single site mutation. In crosses between the mutant and wild type testers, progeny segregated 1 : 1 for parental types only and all hygromycin B-resistant progeny

produced lesions similar to the mutant parent; all hygromycin B-sensitive progeny produced wild type lesions (Figure 14), indicating that a tagged mutation is responsible for the reduced pathogenicity of the mutant.

EXAMPLE 3 -- *Cloning and sequencing of DNA flanking the REMI vector insertion site.*

A total of 11.3 kb of genomic DNA surrounding the insertion site was cloned and completely sequenced (Figure 15). The sequence was derived from seven plasmid clones. The first three (p214B7, p214M1 and p214S1) were recovered from the tagged site in mutant R.C4.2696 and cover about 60 % (6.6 kb) of the entire region. The rest (p118B14, p118BC4, p9P2 and p12H6) were recovered from transformants generated using the chromosome walking strategy. DNA to the left of the insertion site (3.4 kb) was cloned on p214B7; DNA on the right (7.9 kb) was cloned on different overlapping plasmids. p9P2 carries the largest amount (4.6 kb) including genomic DNA on p12H6 (Figure 15).

EXAMPLE 4 -- *Identification of CPS1 and TES1 at the sequenced region.*

Analysis of the combined sequences revealed two open reading frames (ORFs). ORF1(5.4 kb) starts 576 bp upstream of the REMI vector insertion site and ends with an in-frame stop codon (TAG) 3029 bp from the end of the sequenced region in the right flank (Figure 15). No "TATA" box-like element is found in the expected position, but five putative "CAAT" boxes are located upstream of the start codon (ATG), three of them are in the range found in most filamentous fungal promoters (60-200 bp) (Gurr et al., 1987, which is hereby incorporated by reference). Sequence around ATG of ORF1 (CACCATGCT (SEQ. ID. No. 38)) is similar to the fungal consensus (CACCATGGC (SEQ. ID. No. 39)). Although there are several ATGs found upstream, they are less likely to be used as a start codon because the surrounding sequences lack similarity to the consensus. Three putative introns are identified by their conserved 5' and 3' border sequences and potential branch sites (Table 3). Splicing these introns eliminated stop codons which would otherwise interrupt the 5.4 kb open reading frame. Three introns have similar size (45-53 bp respectively) which is in the range of intron size determined from most fungal genes. A putative

polyadenylation signal (ATAA) is found 223 bp downstream of the translation termination site. The G+C content of ORF1 is 51.5%, which is similar to most *Cochliobolus* genes (Turgeon et al., "Cloning and Analysis of the Mating Type Genes from *Cochliobolus heterostrophus*," Mol. Gen. Gene., 238(1-2):270-284 (1993); VanWert et al., "Structure of the *Cochliobolus-heterostrophus* Glyceraldehyde-3-Phosphate Dehydrogenase Gene," Curr. Genet., 22(1):29-35 (1992); Yang et al., "A Polyketide Synthase is Required for Fungal Virulence and Production of the Polyketide T-Toxin," Plant Cell, 8(11):2139-2150 (1996); Rose et al., "A Decarboxylase Required for Polyketide Toxin Production and High Virulence by *Cochliobolus heterostrophus*," 8th Int. Symp. Mol. Plant-Microbe Int., Knoxville, p. J-49 (1996), which are hereby incorporated by reference). Interestingly, ORF1 is flanked by two regions of G+C rich DNA. The first (1.4 kb, 60.7% G+C) is found between ORF1 and ORF2; the second (1.2 kb, 60.3% G+C) is found 1.8 kb downstream of the stop codon of ORF1 (Figure 16). Database searches using the translated protein sequence of ORF1 revealed high similarity to SafB, one of the multifunctional enzymes catalyzing the biosynthesis of the cyclic peptide antibiotic saframycin Mx1 produced by the bacterium *Myxococcus xanthus* (Pospiech et al., "Two Multifunctional Peptide Synthetases and an O-methyltransferase are Involved in the Biosynthesis of the DNA-Binding Antibiotic and Antitumour Agent Saframycin Mx1 from *Myxococcus xanthus*," Microbiology, 142(4):741-746 (1996), which is hereby incorporated by reference). The entire nucleotide sequence of ORF1 (designated *CPS1*) is given in Figure 17.

Table 3. Characteristics of putative introns in *CPS1* and *TES1*

Gene	Intron	Size(bp)	Location	5'Border	3'Border	Branch Site
CPS1	I	45	3060-3105	GTAAGT	TAG	GTCTAAC
	II	51	4532-4582	GTAAGT	CAG	TGCTAAC
	III	53	5187-5239	GTACGT	CAG	TACTAAC
TES1	I	49	528-566	GTAAGT	TAG	CCTTAAG
Consensus				$GTA^{A}/_{C}GT$	$T/_{C}AG$	YNCTAAC*

* Y = Pyrimidine (T or C); N = purine or pyrimidine.

ORF2 starts about 1.6 kb upstream of the start codon of *CPS1* and is transcribed in the opposite direction (Figure 15). No "TATA" box-like element and CAAT box are found; instead, an AT-rich sequence "AAAACTAT" (SEQ. ID. No. 40) is located 11 bp upstream of the start codon ATG and a CT motif is found in the -30 region, which is characteristic of a number of fungal genes that lack a CAAT box in their promoter region (Gurr et al., "The Structure and Organization of Nuclear Genes of Filamentous Fungi," in Kinghorn, ed., Gene Structure in Eukaryotic Microbes, Vol. 22, published by the Society for General Microbiology, Oxford, England:IRL Press, pp. 93-140 (1987), which is hereby incorporated by reference). The sequence around ATG matches perfectly fungal gene consensus. A putative intron (50 bp) is found in the middle of ORF2 with conserved 5' and 3' border sequences and a potential branch site (Table 3). A putative polyadenylation signal (AAATA) is found 189 bp downstream of the translation stop codon TGA. The G+C content of ORF2 is 55.5%, which is slightly higher than the normal range because the 5' end of ORF2 is located in the region of G+C rich DNA upstream of ORF1 (Figure 16). Database search revealed that ORF2 encodes a protein with high similarity to *Homo sapiens* thioesterase II (hTE, Liu et al., "Binding of HIV-1 Nef to a Novel Thioesterase Enzyme Correlates with Nef-Mediated CD4 Down-Regulation," J. Biol. Chem., 272(21:13779-13785 (1997), which is hereby incorporated by reference) and *E. coli* thioesterase II encoded by the *tesB* gene (Naggert et al., "Cloning, Sequencing and Characterization of *Escherichia-coli* Thioesterase II," J. Biol. Chem., 266(17)11044-11050 (1991), which is hereby incorporated by reference). The nucleotide sequence of ORF2 (designated *TES1*) is given in Figure 18.

EXAMPLE 5 -- ***Modular structure of CPS1.***

Predicted CPS1 protein (1743 amino acids, M_r 193235) contains two structurally similar modules, both of which are similar to SafB1, the first module of saframycin synthetase B (overall 25% identity; 50% similarity) and

have apparent amino-acid-activating and thiolation domains but lack methyltransferase activity, thus appearing to be typical type I modules (Figure 19). The number of amino acids in each module is different: the first module (CPS1A) consists of 574 amino acids (from the first residue of core 1 to the last residue of core 6), which is larger than most type I modules; the second module (CPS1B) has 530 amino acids, which is average. The distance between the two modules is 193 amino acids, much shorter than most peptide synthetases (500-600 aa), but this distance is not highly conserved, i.e., an opposite variation is found in HC-toxin synthetase and cyclosporine synthetase, both of which have about 1,000 aa between the first and second amino-acid-activating module (Figure 20F).

Amino acid alignment of the two modules of CPS1 to SafB1 indicated that these modules are highly similar to each other in both overall amino acid composition and conserved motif sequences as defined by Stachelhaus and Marahiel (Stachelhaus et al., "Modular Structure of Peptide Synthetases Revealed by Dissection of the Multifunctional Enzyme GrsA," J. Biol. Chem., 270(11):6163-6169 (1995); Marahiel, "Protein Templates for the Biosynthesis of Peptide Antibiotics," Chem. Biol., 4(8):561-567 (1997), which are hereby incorporated by reference). When aligned to other bacterial or fungal peptide synthetases, CPS1 only showed local similarity to cyclosporine synthetase (Weber et al., "The Peptide Synthetase Catalyzing Cyclosporine Production in *Tolypocladium niveum* is Encoded by a Giant 45.8-Kilobase Open Reading Frame," Current Genetics, 26(2):120-125 (1994), which is hereby incorporated by reference) and tyrocidine synthetase A (Mootz et al., "The Tyrocidine Biosynthesis Operon of Acillus brevis: Complete Nucleotide Sequence and Biochemical Characterization of Functional Internal Adenylation Domains," J. Bacteriol., 179(21):6843-6850 (1997), which is hereby incorporated by reference), but when the amino acids in motif regions were aligned, a overall conservation was observed. Both CPS1A and CPS1B have all five core sequences in the amino-acid-activating domain (Figure 20A-E). Cores 3 and 4 are well conserved except for the replacement of an aspartic acid residue of core 4 by a leucine in CPS1A. Cores 1, 2 and 5 show weak conservation, but similar variations are also seen in SafB1. A thiolation domain is found in both modules, which contains a highly conserved motif (core 6, Figure 20F). The serine residue in this motif has

been shown to be the active site for 4'-phosphopantetheine attachment (Schlumbohm et al, "An Active Serine is Involved in Covalent Substrate Amino Acid Binding at Each Reaction Center of Gramicidin S Synthetase," J. Biol. Chem., 266(34):23135-23141 (1991); Stein et al., "Detection of 4'-Phosphopantetheine at the Thioester Binding Site for L-Valine of GramicidinS Synthetase 2," FEBS Lett., 340(1-2):39-44 (1994), which are hereby incorporated by reference). The distances between the six core sequences in the two modules are also largely conserved. Two exceptions are found in the first module, which has 312 aa between cores 2 and 3, larger than normal (150-200); 61 between cores 5 and 6, only half of that of most peptide synthetases. SafB1 also shows distance variations at these two interval regions (Figure 20B and E). In addition to amino-acid-activating and thiolation domains, CPS1 also has an integrated thioesterase domain (TE) in the carboxy-terminal end of CPS1B (Figure 19). A signature sequence GXSXG, which is highly conserved in animal fatty acid thioesterase type II enzymes and several peptide synthetases, is found in this domain (Figure 21).

EXAMPLE 6 -- ***Sequence homology analysis of TES1 protein.***

The predicted TES1 protein consists of 367 amino acids (M_r 41013). Amino acid alignment of TES1 to hTE, TESB and *Mycobacterium tuberculosis* TESB homolog (Philipp et al., "An Integrated Map of the Genome of the Tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and Comparison with *Mycobacterium leprae*," Proc. Natl. Acad. Sci. USA, 93(7):3132-3137 (1996), which is hereby incorporated by reference) showed that these proteins have an overall 40% identity and 60% similarity. A highly conserved VHS motif (putative active site) is found in the *C*-terminal region of TES1 at a conserved position (Figure 22). All these thioesterases have no sequence similarity with the previously identified animal type I or type II thioesterases known to be involved in the chain termination of fatty acid synthesis (Naggert et al., "Cloning, Sequencing and Characterization of *Escherichia-coli* Thioesterase II," J. Biol. Chem., 266(17):11044-11050 (1991), which is hereby incorporated by reference). Interestingly, TES1 has more homology to hTE than to two bacterial genes, suggesting that both proteins belong to a new family of eucaryotic thioesterases.

EXAMPLE 7 -- *Targeted disruption of CPS1.*

Disruption of either CPS1A or CPS1B restored the original mutant phenotype. Ten transformants from each of four individual disruption experiments using different constructs, including the plasmid recovered from the REMI insertion site in the mutant (p214B7) and three vectors for chromosome walking (p214SNP, p118BSP and p118BCS) were purified and assayed on N-cytoplasm corn. All transformants showed the same small lesion phenotype as that of the original REMI mutant. Gel blot analysis confirmed that all transformants showing the mutant phenotype resulted from homologous integration of the transforming vector that disrupted the wild type *CPS1* (Figures 24-26). No transformants showing the wild type phenotype were obtained, presumably because of the large genomic DNA fragments (over 800 bp in all disruption experiments) on the transforming vector that resulted from high efficiency of homologous recombination and the low chance to recover transformants with ectopic integration.

EXAMPLE 8 -- *Methods and Materials for Examples 9-10*

Strains, growth conditions and transformation. Strains of *Cochliobolus* species and relatives used for genomic DNA hybridization are listed in Table 4. The strain HvW, a victorin-producing isolate of *C. victoriae* was recovered from storage and grown on CMX medium (Turgeon et al., "Transformation of the Fungal Maize Pathogen *Cochliobolus heterostrophus* Using the *Aspergillus nidulans amdS* Gene," Mol. Gen. Genet., 201:450-453 (1985), which is hereby incorporated by reference) for conidiation or on oat meal agar medium (Churchill et al., "Victorin-Deficient REMI Mutants of *Cochliobolus victoriae* Demonstrate a Requirement for Victorin in Pathogenesis," Fungal Genet. Newsl., 42A:41 (1995), which is hereby incorporated by reference) for victorin detection at 24°C under warm white lights (Sylvania Inc., Danvers, MA). Transformation was done using the *C. heterostrophus* procedure (Turgeon et al., "Cloning and Analysis of the Mating Type Genes from *Cochliobolus heterostrophus*," Mol. Gen. Gene., 238(1-2):270-284 (1993), which is hereby incorporated by reference).

Table 4. Detection of *CPS1* homologs in *Cochliobolus* spp and relatives

Strain[a]	Host[b]	Hybridization *Eco*RI digest[c]	Hybridization *Hin*dIII digest[d]	Hybridization *Bgl*II digest[e]
C. heterostrophus	Corn			
race T (C4)	(*Turf-13*)	+	5.2 3.2	4.2
race O (C5)		+	5.2 3.2	4.2
C. carbonum	Corn[1]			
race 1 (26R13)	(*hm1hm1*)	+	6.6	5.0
race 2 (YugY)		N	6.6	5.0
race 3 (BZ1703)*		N	6.6	5.0
C. victoriae (HvW)	Oats (*Vb*)	+	N	5.0
C. sativus (A20)	Grasses[2]	+	3.0	N
C. specifer (D5-7)	Grasses[2]	+	N	N
C. homomorphus (ATCC 13409)	Unknown	N	5.8	N
C. dactyloctenii (7938-9)	Unknown	N	5.9	N
S. turcica (NK2)	Sorghum and maize[3]	+	N	N
S. rostrata (32197)	Weeds and bamboo[4]	+	2.8	N
B. sacchari	Sugarcane[5]			
(764-1)		+	5.4 2.5	N
(1249-10)		N	5.4 2.5	N

a. *C.* = *Cochliobolus*. *S.* = *Setosphaeria*. *B.* = *Bioplaris*. The name of isolates (or lab strains) of each species are given in parentheses and those known to produce host-specific toxins are underlined. * Provided by Tsukiboshi Takao (Japan) and the isolate could be either BZ1209 or BZ1703.

b. Genotype susceptible to the host-specific toxin-producing isolate is given in parentheses. References for hosts of those species not mentioned in the previous chapters are as follows: 1: Welz et al., "Phenotypic Variation and Parasitic Fitness of Races of *Cochliobolus-carbonum* on Corn in North Carolina," Phytopathology, 83(6)593-601 (1993); Leonard et al., "Genetic Diversity in Field Populations of *Cochliobolus-carbonum* on Corn in North Carolina USA," Phytopathology, 80(11):1154-1159 (1990) (for races 2 and 3 only). 2: Domsch et al., "Compendium of Soil Fungi, Vol. 1," New York, New York:Academic Press, pp. 216-222 (1980). 3: David et al., "Fungi on Plants and Plant Products," St. Paul, Minnesota:APS Press, p. 635 (1989); Thakur et al., "Characterization of a New Race of *Exserohilum-turcicum* Virulent on Corn With Resistance Gene HTN," Plant Dis., 73(2):151-155 (1989). 4: Rao et al., "New Fungal Diseases on Some Weeds," Indian Bot. Rep., 6(1):38 (1987); Bhat et al., "Unrecorded

Pathogen on Bamboo Causing Blight in India," Curr. SCI. (BANGALORE), 58(20):1148-1149 (1989). 5: Yoder, "Toxins in Pathogenesis," Ann. Rev. Phytopathol., 18:103-129 (1980).

c. Genomic DNAs (from a previously prepared gel blot filter, Rose et al., "A Decarboxylase Required for Polyketide Toxin Production and High Virulence by *Cochliobolus heterostrophus*," 8th Int. Symp. Mol. Plant-Microbe Int., Knoxville, p. J-49 (1996) were probed with the 3.4 kb *CPS1* fragment cloned on p214B7 (Figure 2). "+" indicates a strong hybridization signal. All species hybridized to a large fragment (~23 kb).

d. Genomic DNAs selected from a lab collection were probed with the *CPS1* 3.2 kb fragment cloned on p214S1 (Figure 2). The size of fragments that hybridized to the probe is given in kb. The intensities of hybridization signals were similar to each other. N = not done.

e. Genomic DNAs were probed with the same *CPS1* fragment as in c. Gel blot is shown in Figure 26.

DNA manipulations and targeted disruption of the CPS1 homolog in C. victoriae. Genomic DNAs for probing were prepared according to Yoder ("*Cochliobolus heterostrophus*, Cause of Southern Corn Leaf Blight," in Sidhu, ed., Genetics of Plant Pathogenic Fungi, Vol. 6, San Diego, California:Academic Press, pp. 93-112 (1988), which is hereby incorporated by reference), or selected from a lab DNA collection (stored at 4°C). A gel blot filter bearing known genomic DNAs was also probed. Plasmid DNA preparation, restriction enzyme digestions, gel electrophoresis, gel blot analysis were done using standard protocols (Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, New York:Cold Spring Harbor Laboratory Press (1989), which is hereby incorporated by reference). For probing, *CPS1* fragments of *C. heterostrophus* cloned on p214B7 (3.4 kb left flank) and p214S1 (3.2 kb right flank) (Figure 2) were prepared by restriction enzyme digestion of the plasmid DNAs followed by purification using the QIAquick Gel Extraction Kit (QIAGEN Inc., Chatsworth, CA). The plasmid p118B14, which carries the 2.2 kb *Bgl* II fragment of *CPS1* interrupted by the *hygB* cassette (see Figure 5) was linearized with *Bgl*II and introduced into HvW genome. Transformants were purified by isolation of single conidia and genomic DNAs were digested with *Bgl*II and probed with the *CPS1* 3.2 kb fragment.

Bioassays. Pathogenicity was determined by an oat plant assay. Fungal strains were grown on individual oat meal agar medium plates (60 X 15 mm) containing hygromycin B (60 ug/ml) for 10 days at 24°C under lights.

Conidia were scraped from the plates and suspended in 6 ml of sterilized distilled water. One ml of conidial suspension of each strain was mixed with 60 seeds of susceptible or resistant oats. Inoculated seeds were planted in 4" X 6" pots and seedlings were allowed to grow for two weeks. Seed germination rate and symptom development were recorded at different stages (4, 6, 8 and 24 days after inoculation). Detection of victorin production using HPLC analysis was done by Alice Churchill in Dr. Vladimir Macko's lab at Boyce Thompson Institute for Plant Research.

EXAMPLE 9 -- ***Detection of CPS1 homologs.***

Genomic DNAs of 12 isolates (or lab strains) of 9 fungal species hybridized to *CPS1* (Table 4). All 6 *Cochliobolus* species, including 4 known plant pathogens (*C. carbonum*, *C. victoriae*, *C. sativus* and *C. spicifer*) and 2 species with unknown hosts (*C. homomorphus* and *C. dactyloctenii*) gave hybridization signals of the same intensity as that of *C. heterostrophus CPS1* fragments (Figure 26, only *C. carbonum* and *C. victoriae* are shown). Two phytopathogenic *Setosphaeria* species and *Bioplaris sacchari*, a sugarcane pathogen gave a similar hybridization intensity.

CPS1 homologs appear to be polymorphic among different species, i.e., all species gave one or two unique bands when *Bgl*II or *Hin*dIII digested genomic DNAs were probed (except for *C. victoriae*, which showed the same hybridization pattern as *C. carbonum*) (Table 4 and Figure 26). Interestingly, *Eco*RI digested genomic DNAs of the same species did not show polymorphisms; all species hybridized to a large fragment (~23 kb, Table 4), indicating the absence of an *Eco*RI site in all *CPS1* homologs as in the *C. heterostrophus* gene. In *C. heterostrophus*, a > 12 kb of genomic region which includes *CPS1* (5.4 kb), *TES1* (1.1 kb) and sequence downstream of the 3' end of *CPS1* has no *Eco*RI sites. In contrast to species-dependent polymorphisms, *CPS1* homologs appear to be highly conserved among different isolates of the same species. Both *C. heterostrophus* race T and race O hybridized to the same 4.2 kb *Bgl*II fragment (or 5.2 and 3.2 kb *Hin*dIII fragments); all three *C. carbonum* races hybridized to the same 5.0 kb *Bgl*II fragment (or 6.6 kb *Hin*dIII fragment) (Table 4, Figure 26) and

B. sacchari isolates 764-1 and 1249-10 hybridized to the same *Hin*dIII fragments (5.4 and 2.5 kb) (Table 4).

EXAMPLE 10 -- ***Targeted disruption of CPS1 homolog in C. victoriae.***

Twenty transformants were obtained from transformation of the victorin-producing isolate HvW with *Bgl*II-linearized plasmid p118B14 (Figure 5). Six transformants were purified and assayed for both victorin production and pathogenicity to susceptible oat plants. All transformants produced wild type levels of victorin as determined by HPLC analysis (Figure 27), but four of them (Tx7, Tx2, Tx5 and Tx8) showed dramatically reduced virulence in the plant assay. The seed germination rate on the eighth day after inoculation is only 13-25% for wild type and two transformants (Tx9 and Tx4), but 45-63% for the other four transformants. On day 24 after inoculation, all plants emerged from the seeds inoculated with wild type, Tx9 or Tx4 were killed but most (29-63%) from the seeds inoculated with Tx2, Tx7, Tx5 or Tx8 still survived (Table 5, Figure 28A). Gel blot analysis confirmed that transformants showing the reduced virulence phenotype resulted from homologous integration of the transforming vector that disrupted the wild type *CPS1* homolog in *C. victoriae* genome; transformants showing the wild type phenotype resulted from ectopic integration events that left the native gene intact (Figure 28B). All transformants remained nonpathogenic to resistant oats, indicating that disruption of the *CPS1* homolog does not affect host specificity of the fungus.

Table 5. Disease development of oat plants inoculated with *C. victoriae* transformants (Tx).

	No. germinated[b]			Germination	No. survivors[d]	
Strain[a]	4	6	8	Rate (%)[c]	24	%
Control-1	28	41	45	75	75	100
Control-2	40	50	50	83	50	100
Control-3	1	7	12	20	0	0
Tx2	8	26	27	45	16	59
Tx4	5	15	15	25	0	0

Tx5	2	24	28	47	8	29
Tx7	14	36	38	63	24	63
Tx8	7	29	29	47	13	47
Tx9	0	3	8	13	0	0

a. Control-1 = uninoculated susceptible oat seeds. Control-2 and Control-3 = resistant and susceptible oat seeds inoculated with wild type *C. victoriae* (isolate HvW), respectively. Six transformants were tested on both resistant and susceptible seeds, but only data for the later are shown (all transformants gave the same results as Control-2 when tested on resistant seeds). Repeat experiments gave similar results (data not shown).
b. Sixty oat seeds were used for each strain. Emerged oat plants were counted 4, 6 and 8 days after inoculation.
c. Calculation based on the data collected on the day 8.
d. Recorded on day 24 after inoculation. The percentage of survivors is based on the number of plants recorded on days 8 and 24.

EXAMPLE 11 -- ***Isolation of CPS1 genes from other plant pathogens***

As dislosed in the previous examples, the *Cochliobolus heterostrophus* gene *CPS1* encodes a peptide synthetase that appears to be a general factor for fungal virulence to their hosts. Thus, *CPS1* has been found to be highly conserved among at least 9 fungal species belonging to 3 genera including the genus *Cochliobolus* and closely related genera *Bioplaris* and *Setosphaeria*; it has been demonstrated to be required for pathogenesis of three different plant pathogens i.e., *C.* heterostrophus race O , race T to corn and *C. victoriae* to oats (Lu, 1998, "Molecular-genetic analysis of general and specific pathogenesis factors in *Cochliobolus heterostrophus*," Ph.D thesis, Cornell University).

To further explore the role of *CPS1* in fungal pathogenesis and its conservation in other fungi, genomic DNAs of additional species in *Cochliobolus* and other closely or distantly related genera were probed with *ChCPS1* by DNA-DNA hybridization (Lu, S.-W., B. G. Turgeon and O. C. Yoder. 1999. "A gene cluster from the corn pathogen *Cochliobolus heterostrophus* required for nonribosomal peptide biosynthesis and general virulence of fungi." Fungal Genetics Conference, March 1999, Pacific Grove, California). Genomic DNAs of 40 filed isolates (or lab strains) representing 34 fungal species belonging to 16 genera hybridized when probed with *ChCPS1* (Figs. 30A-30C).

Fungal genomic DNAs were prepared according to a previously described procedure (Yoder, 1988. "*Cochliobolus heterostrophus*, cause of Southern Corn Leaf Blight". Genetics of Plant Pathogenic Fungi. G. S. Sidhu. San Diego, Academic Press. 6: 93-112). Plasmid DNA preparations, restriction enzyme digestions and preparation of DNA gel blots were performed following standard protocols (Sambrook, J., E. F. Fritsch, et al. 1989 "Molecular Cloning: A Laboratory Manual, 2nd Edition". Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press). A 3.2 kb *ChCPS1* fragment (corresponding to ChCPS1 amino acids 173-1208) was obtained by restriction enzyme digestion of a plasmid clone p214S1 (Lu, 1998, Ph.D thesis), (see Figure 2), followed by purification using the QIAquick Gel Extraction Kit (QIAGEN Inc., Chatsworth, CA). The purified *ChCPS1* fragment was labeled with α-[^{32}P]dCTP (Turgeon, B. G., H. Bohlmann, et al, 1993. "Cloning and analysis of the mating type genes from *Cochliobolus heterostrophus*." Mol. Gen. Genet. 238: 270-284.). DNA-DNA hybridization was carried out at 62° C in 6 X SSC, 0.05 X BLOTTO (Sambrook et al., 1989). Filters were washed in 2 X SSC, 0.1% SDS at 62° C for 60 minutes.

All 16 *Cochliobolus* species, including the known plant pathogens *C. carbonum*, *C. victoriae*, *C.* miyabeanus, *C. sativus* and *C. spicifer*, and five genera closely related to *Cochliobolus*, i.e., *Pyrenophora*, *Setosphaeria*, *Bipolaris*, *Stemphyllium* and *Alternaria* showed hybridization intensities comparable to that of *C. heterostrophus* itself (Fig. 30A).

DNAs of species from nine distantly related genera, including several of economic importance (e. g., *Magnaporthe grisea*, *Fusarium graminearum*, *Gaeumannomyces graminis*) or of medical importance (e. g., *Candida albicans*) hybridized weakly to *CPS1* (FigS. 30B, 30C) whereas no signal was detected in DNA of the basidiomycete *Ustilago maydis*.

Three *CPS1* homolog genes were cloned and characterized. Three of them were cloned from phytopathogenic fungi, including the wheat head scab fungus *Fusarium graminearum* (*FgCPS1*, 6003 bp, SEQ. No. 43), the potato early blight fungus *Alternaria solani*, (*AsCPS1*, 2369bp, SEQ. No. 41) and the barley net blotch fungus *Pyrenophora teres* (*PtCPS1*, 2306 bp, SEQ. No. 45). *FgCPS1* was cloned as a full length gene using both PCR amplification and the plasmid rescue procedure that was preceded by targeted

gene disruption in the genome. *AcCPS1* and *PtCPS1* homologs were partially cloned by direct PCR amplification.

The Polymerase Chain Reaction (PCR) was carried out using degenerate primers designed to conserved regions of *C. heterostrophus CPS1* (*ChCPS1*). Two sets of degenerate primers were designed to amino acids at or close to conserved core sequences of *C. heterostrophus CPS1* (*ChCPS1*). The first pair of primers: 5'TGYTTYATHGCNGGN GTNGTNGCNGTNCC3' (CHFP6, corresponding to positions 493-521 of *ChCPS1*) and 5'YTGYTGNGGNGGNCCNCCNGGRTT3' (CHRP4, 2197-2220 of *ChCPS1*), was used to amplify *CPS1* from *Fusarium graminearum*. The second pair of primers: 5'-AARAARAARGGNCCNACNGAG-3' (FP4CB, corresponding to positions 1531-1550 of *ChCPS1*) and 5'SRYTGNA CCCADATYTCNCC3' (RP2DB, corresponding to positions 3883-3902 of *ChCPS1*), was used to amplify CPS1 from *A. solani* and *Pyrenophora teres*. PCR was carried out in a Perkin Elmer Cetus 9600-thermocycler with fungal (*Fusarium graminearium, Alternaria solani*, and *Pyreophora teres*) genomic DNA as a template.

Reaction mixtures contained about 500 ng of genomic DNA in 100 ul of reaction buffer [1 x *Ex Taq* buffer, 0.2 mM dNTPs, 0.2 uM of each primer and 0.05 U/ml *Takara Ex Taq* (Pan Vera Corporation)]. An initial denaturing step of 95 °C for 3 min. was followed by 30 cycles of 94 °C for 1 min, 47 °C (for *G. zeae*) or 55 °C (for *A. solni* and *Pyrenophora teres*) for 3 min, and 72°C for 3 min. Reactions were cooled to 4 °C after a final extension of 72 °C for 10 min. PCR products were examined (10 ul of each reaction) by agarose-gel (0.75%) electrophoresis.

DNA from isolated clones of the three different fungi was sequenced at the Cornell DNA Sequencing Facility using TaqCycle automated sequencing with DyeDeoxy terminators (Applied Biosystems, Foster City, CA). Primers used for sequencing were designed using Primer Select (DNASTAR) and synthesized by the Cornell Oligonucleotide Synthesis Facility. Sequencing of each plasmid clone was initiated with vector-specific primers or primers designed to previously determined sequences. Sequences were analyzed using MapDraw and MegAlign (DNASTAR) and nucleotide or

protein database searches were performed with the BLAST program (Altschul et al., 1990, 1997).

The *FgCPS1* open reading frame (5125 bp, SEQ. No. 43) has 50% nucleotide identity to *ChCPS1* (SEQ ID No.2) in about 4.4 kb of overlap (Fig. 2). No "TATA" box-like element was found in the 5' untranslated region, but other promoter sequences including two putative "CAAT" boxes and a "CT" motif were located upstream of the start codon (ATG) in *FgCPS1* (Fig. 32). Only one putative intron was found 1508 bp upstream the stop codon (TGA) in contrast to three in *ChCPS1* (Figs. 31A, 31B and 32). A putative polyadenylation signal "AATAA" was located 62 bp downstream of the stop codon (Fig. 32). The predicted FgCPS1 protein (1692 amino acids, M_r 187983 Da, SEQ ID No. 44) has 68% identity, 73% similarity to ChCPS1 (SEQ ID No. 3) in a about 1,500 amino acid overlap (Figs. 31A and 31B) that contains two structurally similar modules highly similar to those of ChCPS1. FgCPS1 has no significant similarity to ChCPS1 at the C-terminus, which is relatively shorter and lacks the thioeterase domain as seen in ChCPS1 (Figs. 31A and 31B). The annotated *FgCPS1* sequence is given in Figure 32.

AsCPS1 (2369 bp, SEQ. No. 41) has 76% nucleotide identity to *ChCPS1* (SEQ ID No.2) in the entire cloned region which contains two conserved introns (Figs. 31A and B). The translated AsCPS1 protein (partial) includes 758 amino acids corresponding to amino acids 511-1269 in ChCPS1 and has up to 93% identity, 95% similarity to ChCPS1 (Fig. 2). The annotated *AsCPS1* sequence is given in Figure 33.

PtCPS1 (2306 bp, SEQ. No. 45) has 78% nucleotide identity to *ChCPS1* (SEQ IDNo.2) in the entire cloned region which contains only one intron (Fig. 2). The translated PtCPS1 protein (partial) includes 758 amino acids corresponding to amino acids 511-1269 in ChCPS1 and has 93% identity, 96% similarity to ChCPS1 (Fig. 2). The annotated *PtCPS1* seqence is given in Figure 34.

EXAMPLE 12 -- ***Targeted disruption of CPS1 homolog in F. graminearum.***

A 2.2 kb XbaI fragment from pUCATPH (Lu, et al, 1994) containing the bacterial hygmycin resistance gene (hygB) driven by the

Aspergillus nidulans trpC promoter was inserted into the *Xba*I site of a PCR clone pFgC8, which carries the 1.0 kb internal fragment of *FgCPS1*, to create pFgC8-hygB. This construct was transformed (in a circular form or linearized with *Hind*III or *Bgl*II) into an isolate (GZ3639) of wild type *F. graminearum*. Twenty transformants were obtained by transformation of a wild type isolate (GZ3696) with circular plasmid pFgC8-hygB. DNA gel blot analysis of eight such transformants confirmed that the *CPS1* homolog in *F. graminearum* was disrupted by a single cross over recombination (Fig. 35A). Two transformants were obtained when the *Bgl*II-digested pFgC8-hygB was used and only one recovered when the plasmid was digested with *Hind*III. Gel blot analysis indicated that all three transformants obtained using linearized plasmid integrated into the genome at a ectopic location that left the wild type *FgCPS1* gene intact (Fig. 35A). For virulence assays, *F. graminearum* strains were grown on PDA (or PDA plus hygromycin B for transformants) plates (100 X 15 mm) for 7-15 days at 24 C under black lights (Sylvania Inc., Danvers, MA) for maximum conidiation. A susceptible spring wheat cultivar, Norm Hard Red (kindly provided by G. Bergstrom, Cornell university) was used. Two months old wheat plants grown in the greenhouse at anthesis (10 plants in one 4" X 6" pot) were sprayed with 10 ml conidial suspensions (10^5 or 10^4 conidia/ml) using a pressurized Preval Spray Gun Power Unit thin layer chromatography sprayer (Alltech Associates, Deerfield, IL) or "injected'' into the spikelets (5-10 ul/ per spikelet) using a Benchmate pipettor (NICHIRYO, Japan). The inoculated plants were incubated in a mist chamber for 48 hours (23° C) and then transferred to a growth chamber (23° C, 14 hours of light). Mutant phenotypes were identified by the comparison of the number of infected or "bleached" spikelets on each head of wheat plants. Symptoms were recorded 7-10 days after inoculation. All *cps1*$^-$ disruptants had reduced virulence on wheat plants in the assays while ectopic transformants caused disease symptoms indistinguishable from those of wild type (Fig. 35B).

Although the invention has been described in detail for the purposes of illustration, it is understood that such detail is solely for that purpose, and variations can be made therein by those skilled in the art without departing

from the spirit and scope of the invention which is defined by the following claims.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule from a plant pathogen encoding a CPS1 synthetase, wherein the nucleic acid molecule comprises a nucleotide sequence which hybridizes under stringent conditions to a nucleic acid molecule having a sequence as set forth in SEQ ID No. 41 or a complement thereof; or wherein the nucleic acid molecule comprises a nucleotide sequence as set forth in SEQ ID No. 41.

2. The isolated nucleic acid molecule according to claim 1 wherein the plant pathogen is selected from the group consisting of *Alternaria solani, Alternaria Alternaia alternatherae, A. alternata, A. amaranthi, A. araliae, A. brassicae, A. brassicicola, A. camelliae, A. cassiae, A. cheiranthi, A. cinerariae, A. gossypii, A. helianthi, A. helianthinficiens, A. mali,* and *A. raphani.*

3. The isolated nucleic acid molecule according to claim 2 wherein the plant pathogen is *Alternaria solani.*

4. An isolated nucleic acid molecule coding for a polypeptide having the amino acid sequence as set forth in SEQ ID No. 42.

5. An isolated nucleic acid molecule comprising the nucleotide sequence as set forth in SEQ ID No. 41.

6. An isolated polypeptide encoded by the nucleic acid molecule of claim 1.

7. The polypeptide according to claim 6, wherein the polypeptide has an amino acid sequence which has at least 75 % similarity to the amino acid sequence of SEQ. ID. No. 42 as determined by a BLAST program with the default parameters.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule from a plant pathogen encoding a CPS1 synthetase, wherein the nucleic acid molecule comprises a nucleotide sequence which hybridizes under stringent conditions to a nucleic acid molecule having a sequence as set forth in SEQ ID No. 41 or a complement thereof; or wherein the nucleic acid molecule comprises a nucleotide sequence as set forth in SEQ ID No. 41.

2. The isolated nucleic acid molecule according to claim 1 wherein the plant pathogen is selected from the group consisting of *Alternaria solani, Alternaria Alternaia alternatherae, A. alternata, A. amaranthi, A. araliae, A. brassicae, A. brassicicola, A. camelliae, A. cassiae, A. cheiranthi, A. cinerariae, A. gossypii, A. helianthi, A. helianthinficiens, A. mali,* and *A. raphani.*

3. The isolated nucleic acid molecule according to claim 2 wherein the plant pathogen is *Alternaria solani.*

4. An isolated nucleic acid molecule coding for a polypeptide having the amino acid sequence as set forth in SEQ ID No. 42.

5. An isolated nucleic acid molecule comprising the nucleotide sequence as set forth in SEQ ID No. 41.

6. An isolated polypeptide encoded by the nucleic acid molecule of claim 1.

7. The polypeptide according to claim 6, wherein the polypeptide has an amino acid sequence which has at least 75 % similarity to the amino acid sequence of SEQ. ID. No. 42 as determined by a BLAST program with the default parameters.

8. The polypeptide according to claim 7 wherein the polypeptide comprises an amino acid sequence as set forth in SEQ ID No. 42.

9. An isolated nucleic acid molecule from a plant pathogen encoding a CPS1 synthetase, wherein the nucleic acid molecule comprises a nucleotide sequence which hybridizes under stringent conditions to a nucleic acid molecule having a sequence as set forth in SEQ ID No. 43 or a complement thereof; or wherein the nucleic acid molecule comprises a nucleotide sequence as set forth in SEQ ID No. 43.

10. The isolated nucleic acid molecule according to claim 1 wherein the plant pathogen is selected from the group consisting of *Fusariumi graminearium, Fusarium avenaceum, F. carpineum, F. chlamydosporum, F.. coccophilum, F. culmorum, F. episphaeria, F. equiseti, F. flocciferum, F. moniliforme, F. oxysporum, F. redolens, F. sambucinum, F. solani, F. subglutinans, F. trichothecioides, F. udum,* or *F. ventricosum.*

11. The isolated nucleic acid molecule according to claim 10 wherein the plant pathogen is *Fusarium graminearium.*

12. An isolated nucleic acid molecule coding for a polypeptide having the amino acid sequence as set forth in SEQ ID No. 44.

13. An isolated nucleic acid molecule comprising the nucleotide sequence as set forth in SEQ ID No. 43.

14. An isolated polypeptide encoded by the nucleic acid of claim 9.

15. The polypeptide according to claim 14, wherein the polypeptide has an amino acid sequence which has at least 75 % similarity to the amino acid sequence of SEQ. ID. No. 44 as determined by a BLAST program with the default parameters.

16. The polypeptide of claim 15 wherein the polypeptide comprises the amino acid sequence as set forth in SEQ ID No. 44.

17. An isolated nucleic acid molecule from a plant pathogen encoding a CPS1 synthetase, wherein the nucleic acid molecule comprises a nucleotide sequence which hybridizes under stringent conditions to a nucleic acid molecule having a sequence as set forth in SEQ ID No. 45 or a complement thereof; or wherein the nucleic acid molecule comprises a nucleotide sequence as set forth in SEQ ID No. 45.

18. The isolated nucleic acid molecule according to claim 1 wherein the plant pathogen is selected from the group consisting of *Pyrenophora teres, Pyrenophora avenae, P. bromi, P. leuceienes, P. phaeocomes, P. schroeteri, P. trichostoma,* or *P. tritici-repentis.*

19. The isolated nucleic acid molecule according to claim 18 wherein the plant pathogen is *Pyrenophera teres.*

20. An isolated nucleic acid molecule coding for a polypeptide having the amino acid sequence as set forth in SEQ ID No. 46.

21. An isolated nucleic acid molecule comprising the nucleotide sequence as set forth in SEQ ID No. 45.

22. An isolated polypeptide encoded by the nucleic acid of claim 17.

23. The polypeptide according to claim 22, wherein the polypeptide has an amino acid sequence which has at least 75 % similarity to the amino acid sequence of SEQ. ID. No. 46 as determined by a BLAST program with the default parameters.

24. The polypeptide of claim 23 wherein the polypeptide comprises the amino acid sequence as set forth in SEQ ID No. 46.

25. A vector comprising the nucleic acid molecule of any one of claims 1-5, 9-13, or 17-21.

26. A vector according to claim 25, wherein the nucleic acid molecule is operably linked to a promoter.

27. A vector according to claim 26, wherein the nucleic acid molecule is in a sense orientation.

28. A vector according to claim 26, wherein the nucleic acid molecule is in an antisense orientation.

29. The vector according to claim 26, wherein the vector is capable of maintaining and expressing the nucleic acid molecule in bacterial cells.

30. The vector according to claim 26, wherein the vector is capable of maintaining and expressing the nucleic acid molecule in plant cells.

31. A host cell transformed with the vector according to claim 25.

32. A host cell according to claim 31, wherein the host is a plant.

33. A host cell according to claim 31, wherein the host is selected from the group consisting of corn, oats, grasses, weeds, bamboo, and sugarcane.

34. A host cell according to claim 31, wherein the host is corn.

35. A plant transformed with the nucleic acid molecule according of any one of claims 1-5, 9-13, or 17-21.

36. An isolated DNA molecule according to claim 1, wherein the nucleic acid molecule has a nucleotide sequence which is 70% identical to the

nucleotide sequence of SEQ. ID. No. 41 as determined by a BLAST program with default parameters.

37. An isolated DNA molecule according to claim 1, wherein the nucleic acid molecule is 80% identical to the nucleotide sequence of SEQ. ID. No. 41 as determined by a BLAST program with default parameters.

38. An isolated DNA molecule according to claim 1, wherein the nucleic acid molecule is 90% identical to the nucleotide sequence of SEQ. ID. No. 41 as determined by a BLAST program with default parameters.

39. An isolated DNA molecule according to claim 1, wherein the nucleic acid molecule has a nucleotide sequence which is 70% identical to the nucleotide sequence of SEQ. ID. No. 43 as determined by a BLAST program with default parameters.

40. An isolated DNA molecule according to claim 1, wherein the nucleic acid molecule is 80% identical to the nucleotide sequence of SEQ. ID. No. 43 as determined by a BLAST program with default parameters.

41. An isolated DNA molecule according to claim 21, wherein the nucleic acid molecule is 90% identical to the nucleotide sequence of SEQ. ID. No. 43 as determined by a BLAST program with default parameters.

42. An isolated DNA molecule according to claim 1, wherein the nucleic acid molecule has a nucleotide sequence which is 70% identical to the nucleotide sequence of SEQ. ID. No. 45 as determined by a BLAST program with default parameters.

43. An isolated DNA molecule according to claim 1, wherein the nucleic acid molecule is 80% identical to the nucleotide sequence of SEQ. ID. No. 45 as determined by a BLAST program with default parameters.

44. An isolated DNA molecule according to claim 21, wherein the nucleic acid molecule is 90% identical to the nucleotide sequence of SEQ. ID. No. 45 as determined by a BLAST program with default parameters.

45. A method for identifying inhibitors of a CPS1 protein , wherein said CPS1 protein is a peptide synthetase of a plant pathogen, said method comprising:

providing a CPS1 protein or polypeptide;

contacting the protein or polypeptide with potential inhibitor compounds;

determining peptide synthetase activity; and

selecting compounds which decrease the peptide synthetase activity.

46. The method of Claim 45 wherein the CPS1 protein is from *Alternaria solani.*

47. The method of Claim 45 wherein the CPS1 protein is from *Fusarium graminearium.*

48. The method of Claim 45 wherein the CPS1 protein is from *Pyrenophora teres.*

49. A method of imparting disease resistance to a plant, said method comprising over-expressing a CPS1 polypeptide in the plant, wherein the polypeptide has protein synthetase activity.

50. A method according to claim 49, wherein the plant is selected from a group consisting of corn, oats, grasses, weeds, sugarcane. barley, wheat, rice, tomato, potato, citrus, malus, rye, cotton, brassica, cabbage, and carrot.

51. The method of claim 49 wherein the CPS1 polypeptide is from *Alternaria solani.*

52. The method of Claim 49 wherein the CPS1 polypeptide is from *Fusarium graminearium*.

53. The method of Claim 45 wherein the CPS1 polyeptide is from *Pyrenophora teres*.

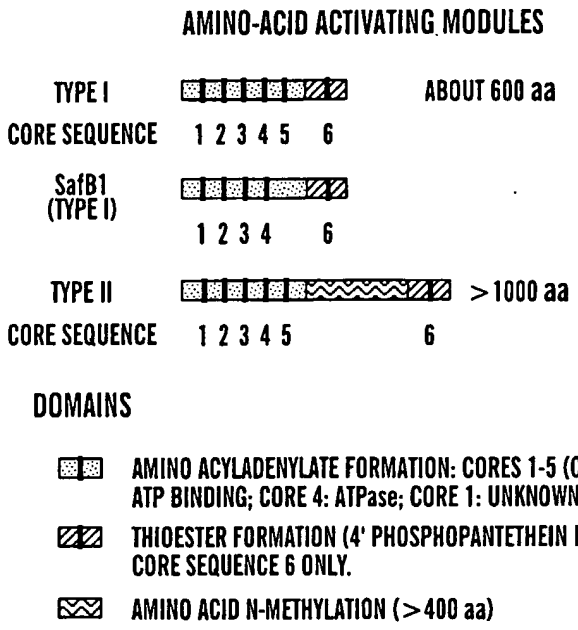

AMINO-ACID ACTIVATING MODULES
TYPE I
ABOUT 600 aa
CORE SEQUENCE 1 2 3 4 5 6
SafB1 (TYPE I)
1 2 3 4 6
TYPE II
>1000 aa
CORE SEQUENCE 1 2 3 4 5 6
DOMAINS
AMINO ACYLADENYLATE FORMATION: CORES 1-5 (CORES 2, 3, AND 4: ATP BINDING; CORE 4: ATPase; CORE 1: UNKNOWN FUNCTION)
THIOESTER FORMATION (4' PHOSPHOPANTETHEIN BINDING), CORE SEQUENCE 6 ONLY.
AMINO ACID N-METHYLATION (>400 aa)

FIG. 1A

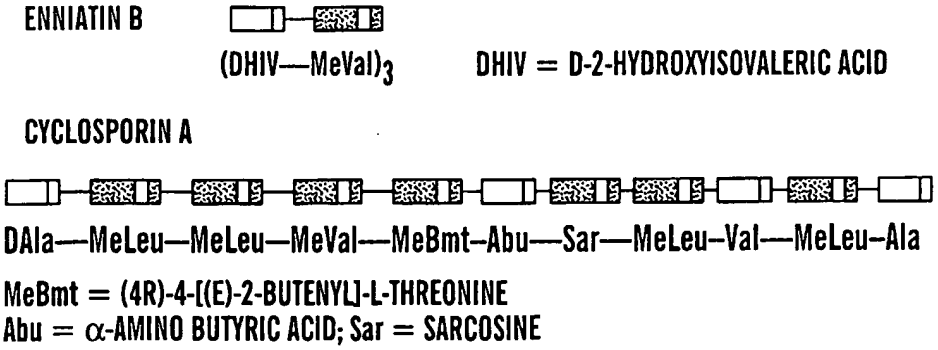

ENNIATIN B
(DHIV—MeVal)3
DHIV = D-2-HYDROXYISOVALERIC ACID
CYCLOSPORIN A
DAla—MeLeu—MeLeu—MeVal—MeBmt–Abu—Sar—MeLeu–Val—MeLeu–Ala
MeBmt = (4R)-4-[(E)-2-BUTENYL]-L-THREONINE
Abu = α-AMINO BUTYRIC ACID; Sar = SARCOSINE

FIG. 1B

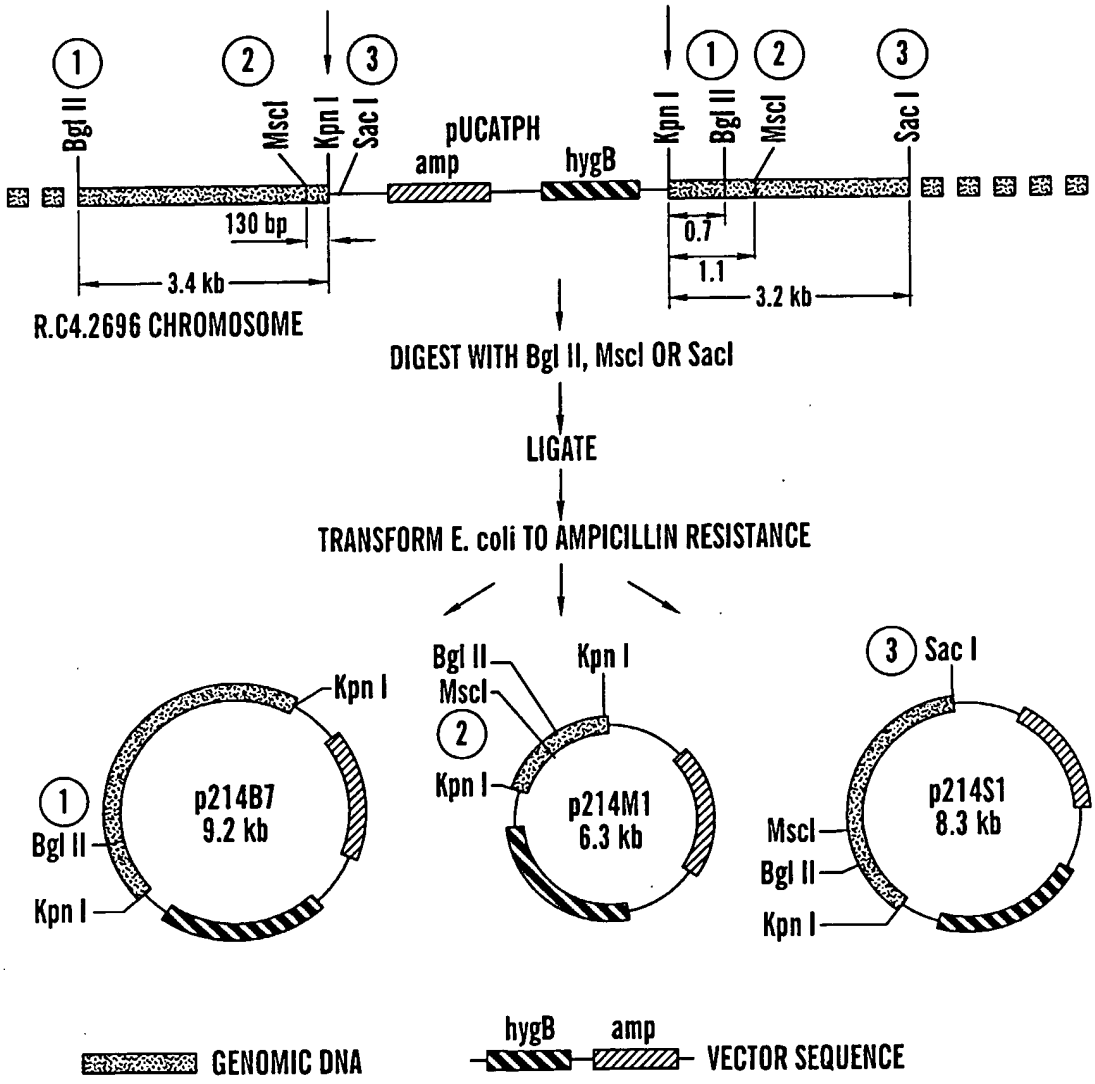

1
2
3
Bgl II
MscI
Kpn I
Sac I
pUCATPH
amp
hygB
Kpn I
Bgl II
MscI
Sac I
130 bp
3.4 kb
0.7
1.1
3.2 kb
R.C4.2696 CHROMOSOME
DIGEST WITH Bgl II, MscI OR SacI
LIGATE
TRANSFORM E. coli TO AMPICILLIN RESISTANCE
p214B7
9.2 kb
p214M1
6.3 kb
p214S1
8.3 kb
GENOMIC DNA
VECTOR SEQUENCE

FIG. 2

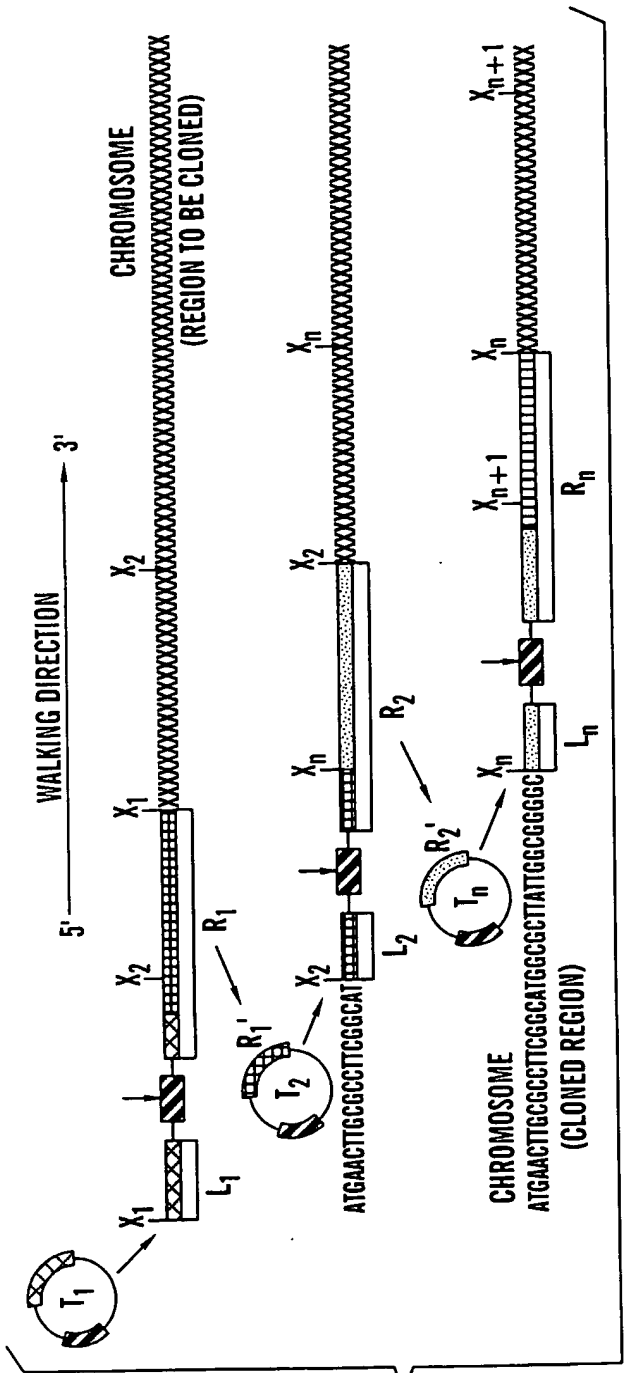

WALKING DIRECTION
5'
3'
CHROMOSOME
(REGION TO BE CLONED)
ATGAACTTGCGCCTTCGGCAT
CHROMOSOME
ATGAACTTGCGCCTTCGGCATGGCGCTTATTGGCGGGGGC
(CLONED REGION)

FIG. 3A

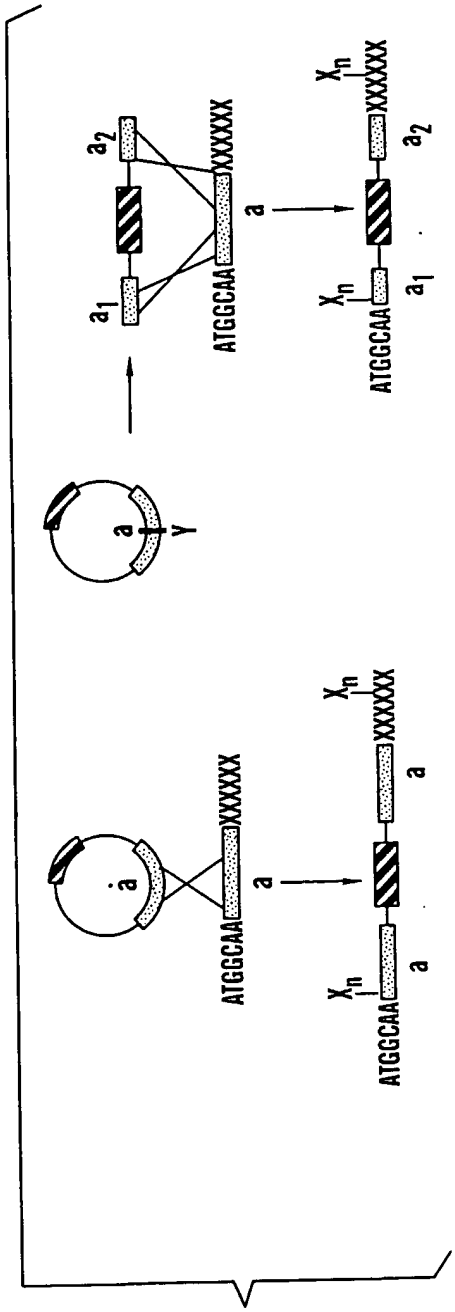

ATGGCAA
ATGGCAA
ATGGCAA
ATGGCAA

FIG. 3B

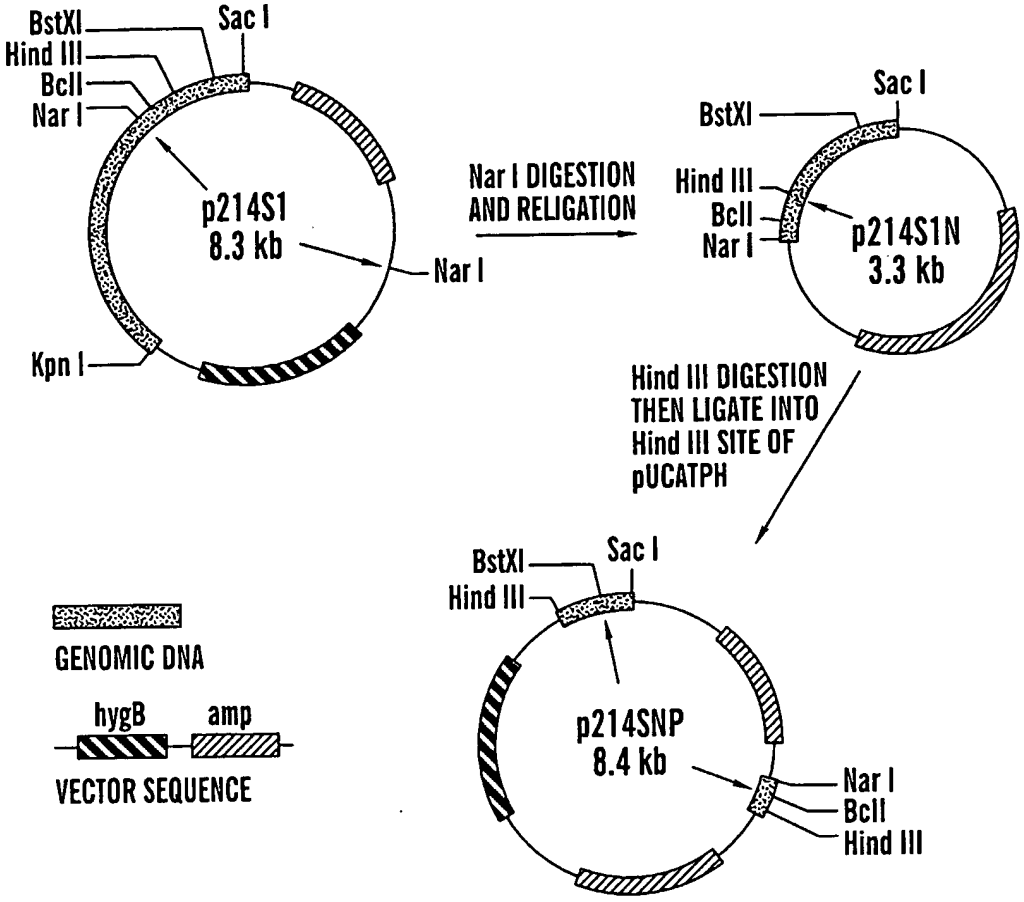

BstXI
Sac I
Hind III
BclI
Nar I
p214S1
8.3 kb
Nar I
Kpn I
Nar I DIGESTION
AND RELIGATION
Sac I
BstXI
Hind III
BclI
Nar I
p214S1N
3.3 kb
Hind III DIGESTION
THEN LIGATE INTO
Hind III SITE OF
pUCATPH
BstXI
Sac I
Hind III
p214SNP
8.4 kb
Nar I
BclI
Hind III
GENOMIC DNA
hygB
amp
VECTOR SEQUENCE

FIG. 4

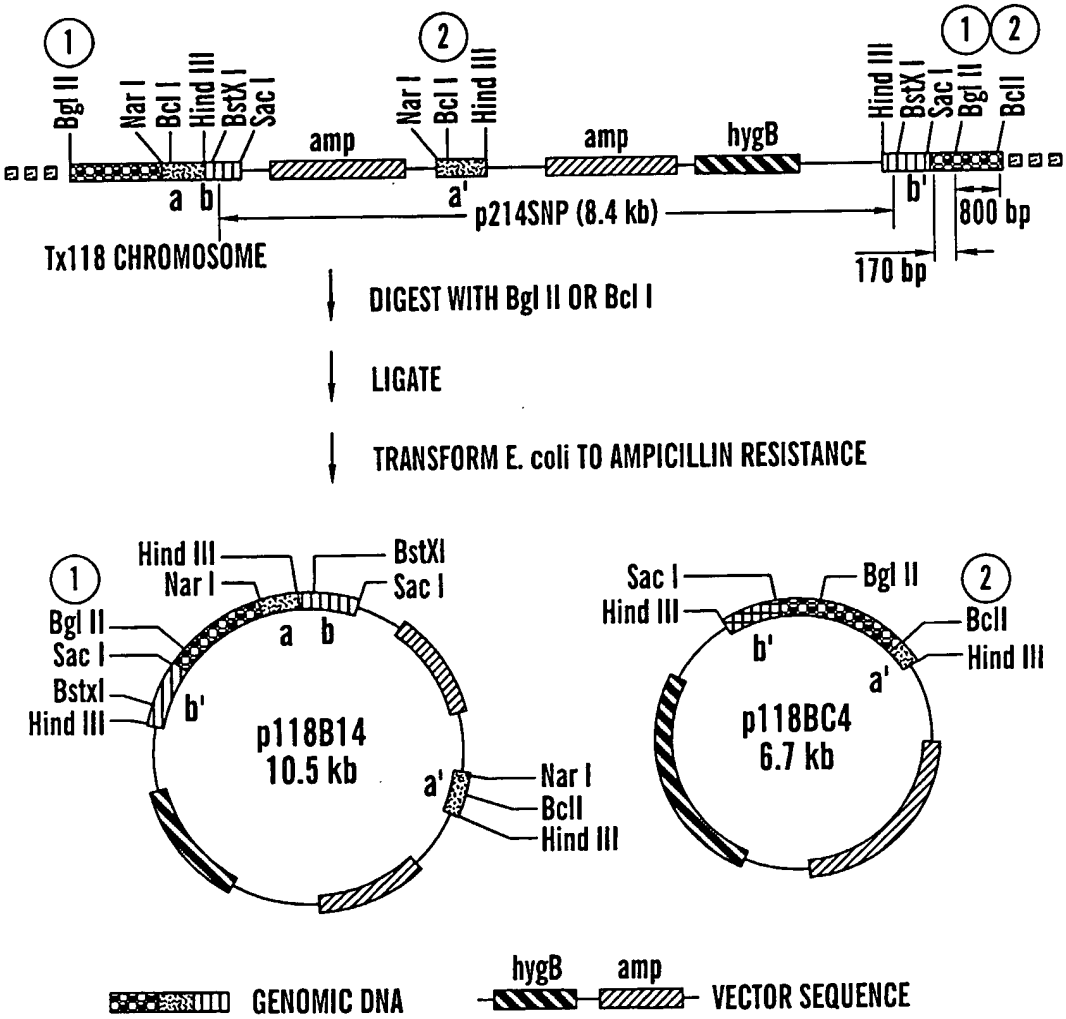

1
Bgl II
Nar I
Bcl I
Hind III
BstX I
Sac I
2
Nar I
Bcl I
Hind III
amp
amp
hygB
1 2
Hind III
BstX I
Sac I
Bgl II
BclI
a b
a'
b'
p214SNP (8.4 kb)
800 bp
Tx118 CHROMOSOME
170 bp
DIGEST WITH Bgl II OR Bcl I
LIGATE
TRANSFORM E. coli TO AMPICILLIN RESISTANCE
1
Hind III
Nar I
BstXI
Sac I
Bgl II
Sac I
BstxI
Hind III
a b
b'
p118B14
10.5 kb
a'
Nar I
BclI
Hind III
Sac I
Hind III
Bgl II
2
BclI
Hind III
b'
a'
p118BC4
6.7 kb
GENOMIC DNA
hygB
amp
VECTOR SEQUENCE

FIG. 5

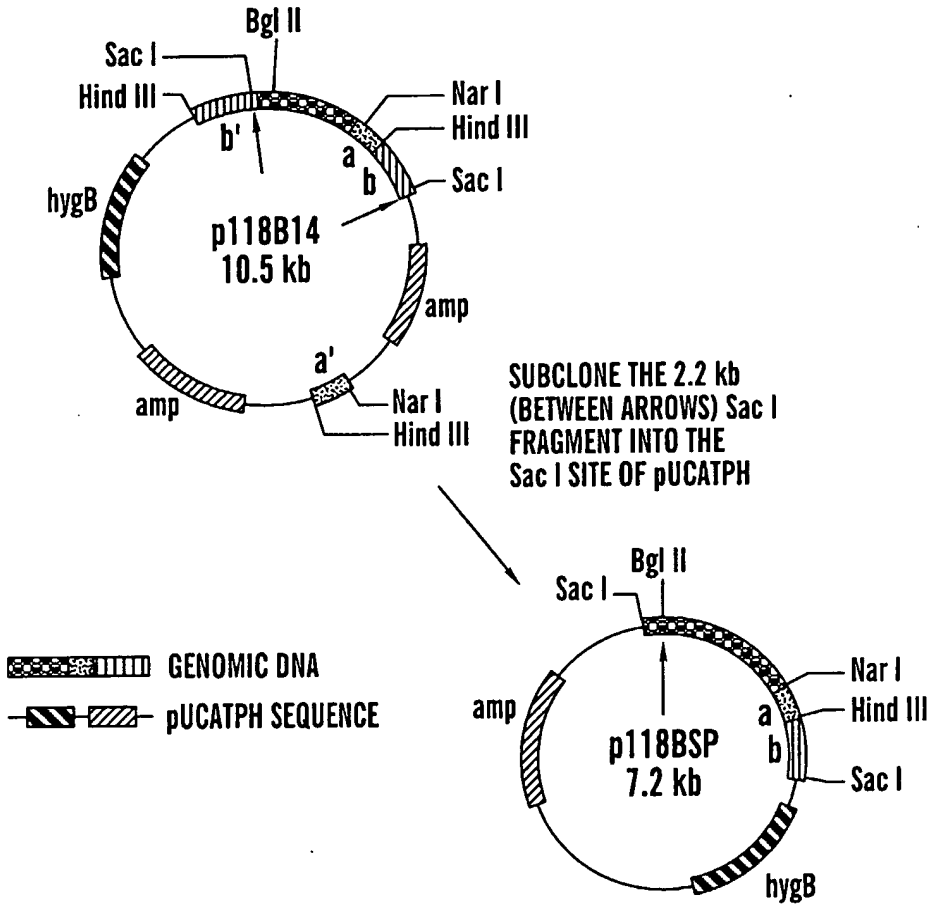

Bgl II
Sac I
Hind III
Nar I
Hind III
Sac I
b'
a
b
hygB
p118B14
10.5 kb
amp
a'
amp
Nar I
Hind III
SUBCLONE THE 2.2 kb
(BETWEEN ARROWS) Sac I
FRAGMENT INTO THE
Sac I SITE OF pUCATPH
GENOMIC DNA
pUCATPH SEQUENCE
Bgl II
Sac I
Nar I
Hind III
amp
a
b
p118BSP
7.2 kb
Sac I
hygB

FIG. 6

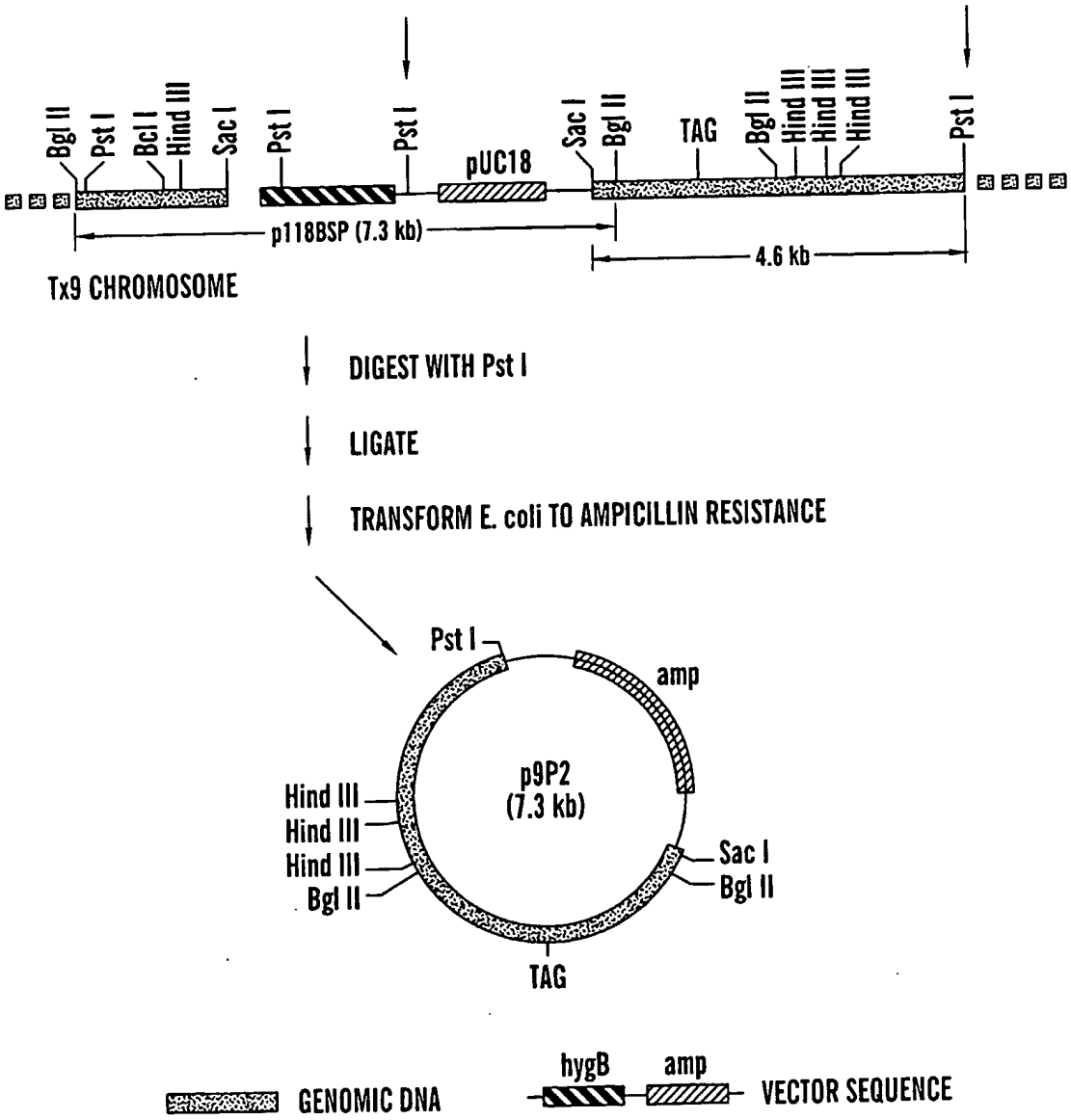

Bgl II
Pst I
Bcl I
Hind III
Sac I
Pst I
Pst I
pUC18
Sac I
Bgl II
TAG
Bgl II
Hind III
Hind III
Hind III
Pst I
p118BSP (7.3 kb)
4.6 kb
Tx9 CHROMOSOME
DIGEST WITH Pst I
LIGATE
TRANSFORM E. coli TO AMPICILLIN RESISTANCE
Pst I
amp
p9P2
(7.3 kb)
Hind III
Hind III
Hind III
Bgl II
Sac I
Bgl II
TAG
GENOMIC DNA
hygB
amp
VECTOR SEQUENCE

FIG. 7

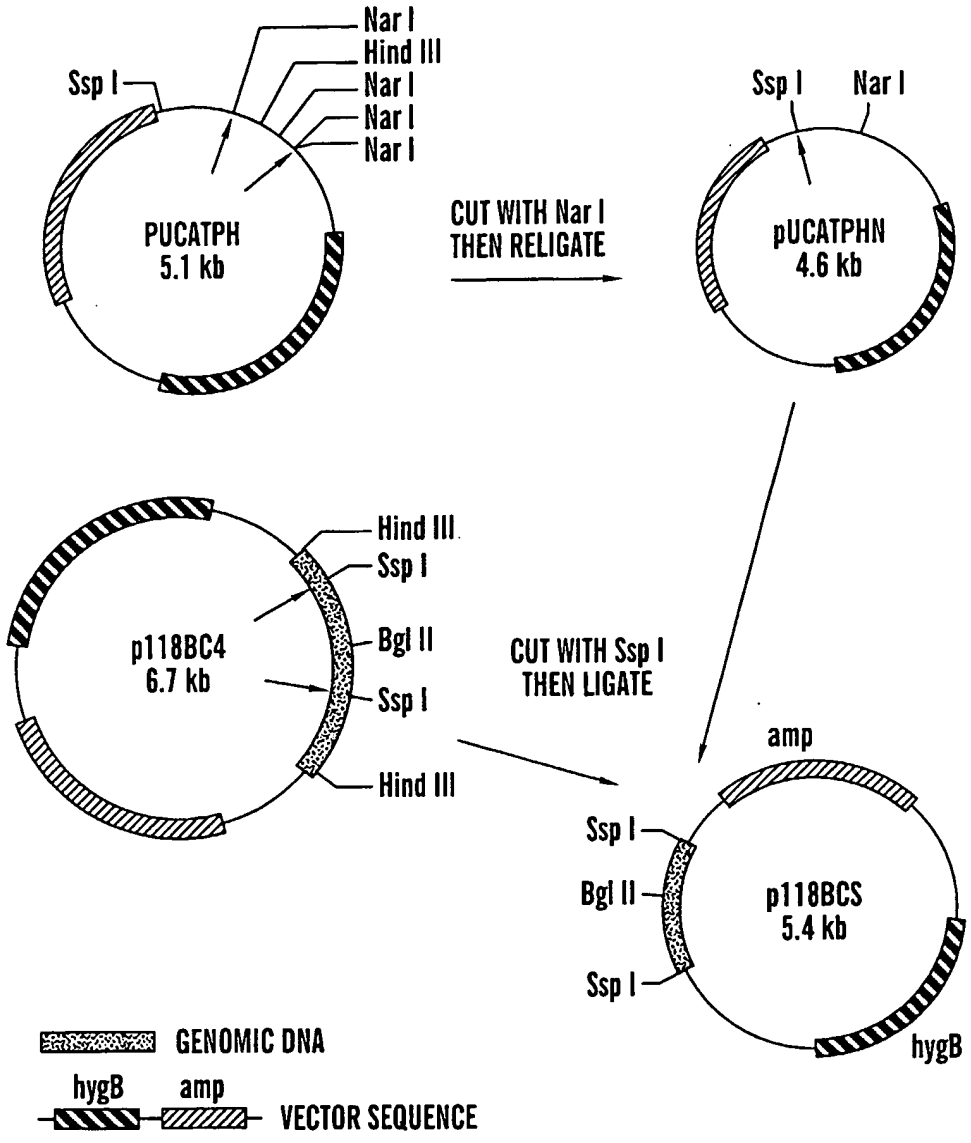

Nar I
Hind III
Nar I
Nar I
Nar I
Ssp I
PUCATPH
5.1 kb
CUT WITH Nar I
THEN RELIGATE
Ssp I
Nar I
pUCATPHN
4.6 kb
Hind III
Ssp I
Bgl II
Ssp I
Hind III
p118BC4
6.7 kb
CUT WITH Ssp I
THEN LIGATE
amp
Ssp I
Bgl II
Ssp I
p118BCS
5.4 kb
hygB
GENOMIC DNA
hygB
amp
VECTOR SEQUENCE

FIG. 8

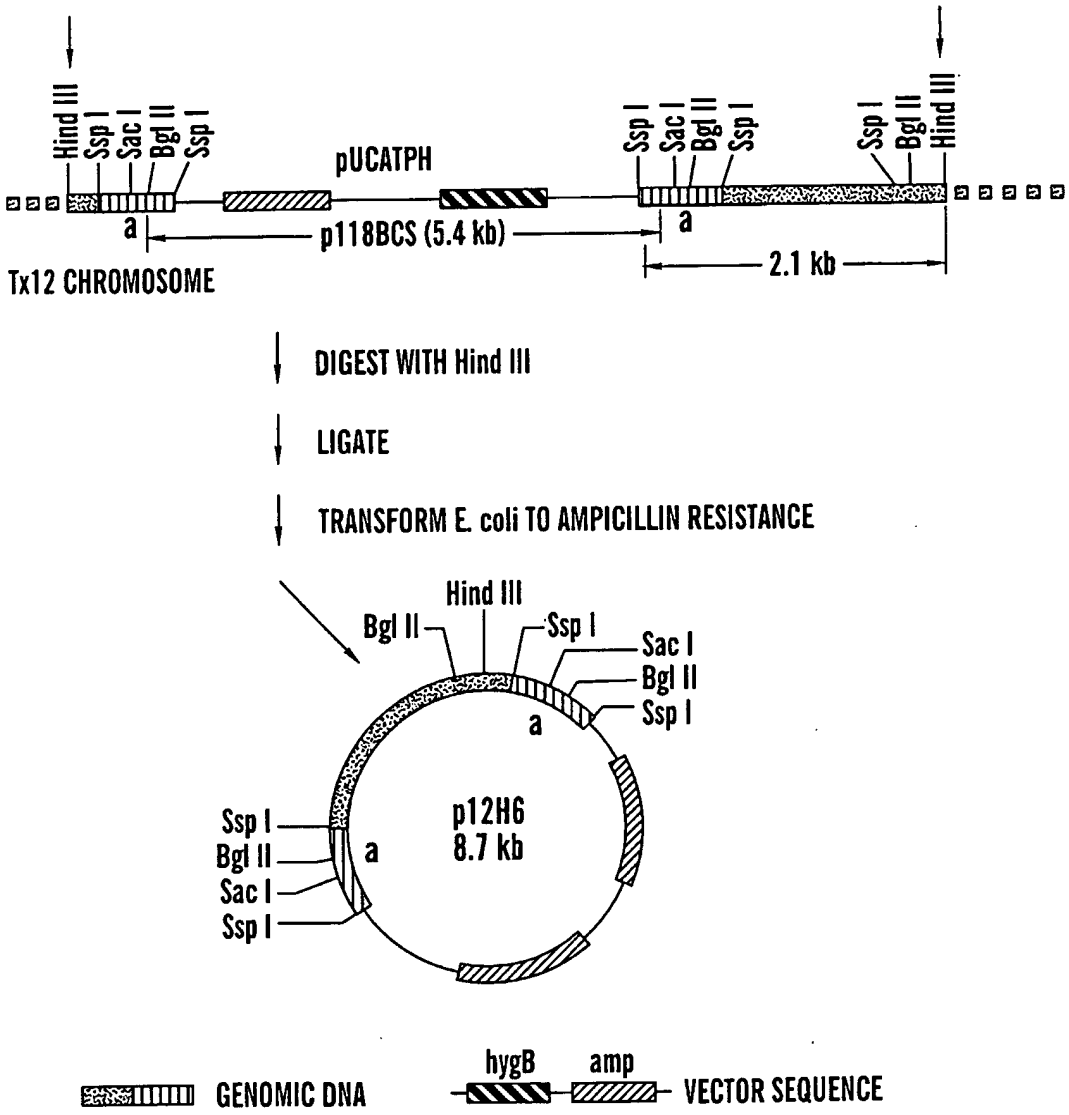

Hind III
Ssp I
Sac I
Bgl II
Ssp I
pUCATPH
Ssp I
Sac I
Bgl II
Ssp I
Ssp I
Bgl II
Hind III
a
p118BCS (5.4 kb)
a
2.1 kb
Tx12 CHROMOSOME
DIGEST WITH Hind III
LIGATE
TRANSFORM E. coli TO AMPICILLIN RESISTANCE
Hind III
Bgl II
Ssp I
Sac I
Bgl II
Ssp I
a
p12H6
8.7 kb
Ssp I
Bgl II
Sac I
Ssp I
a
hygB
amp
GENOMIC DNA
VECTOR SEQUENCE

FIG. 9

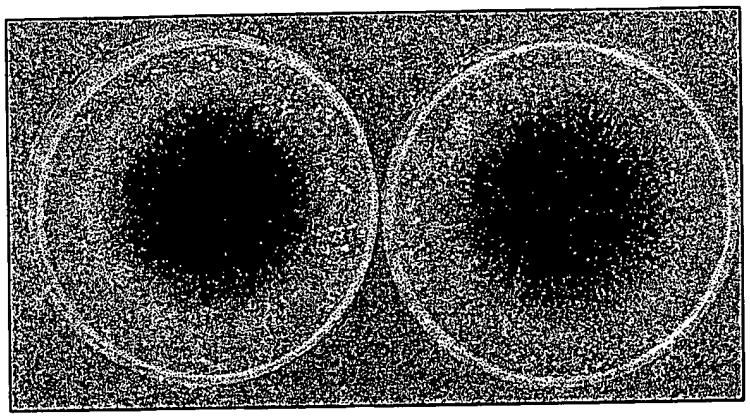


FIG. 10A

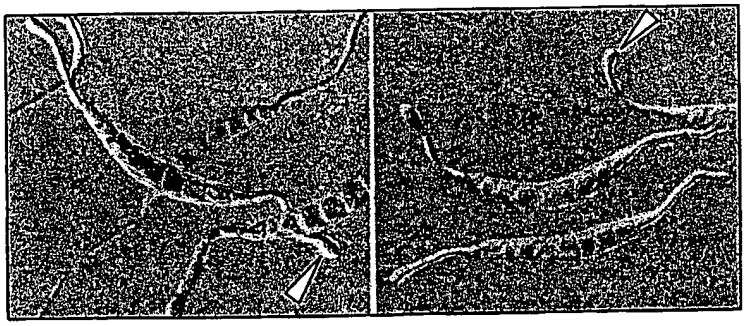


FIG. 10B

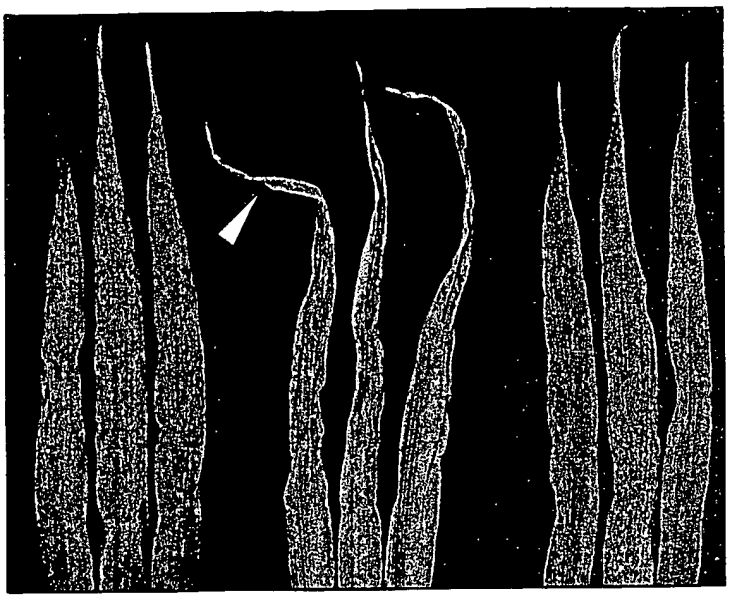


FIG. 11A

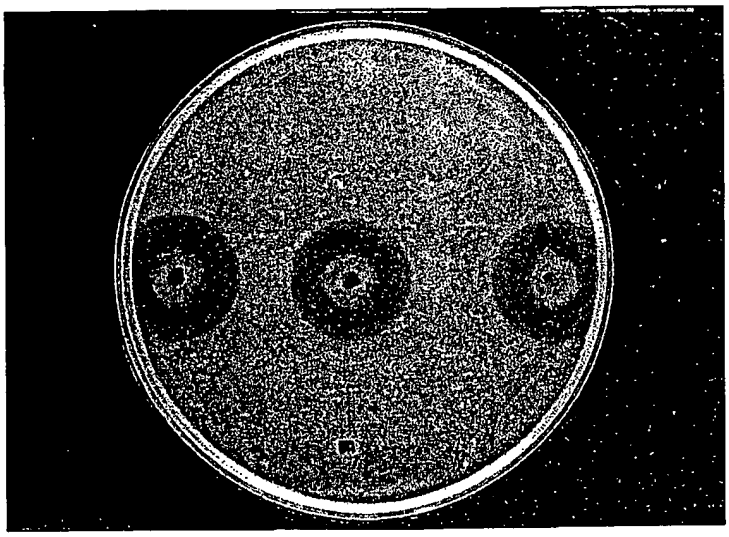


FIG. 11B

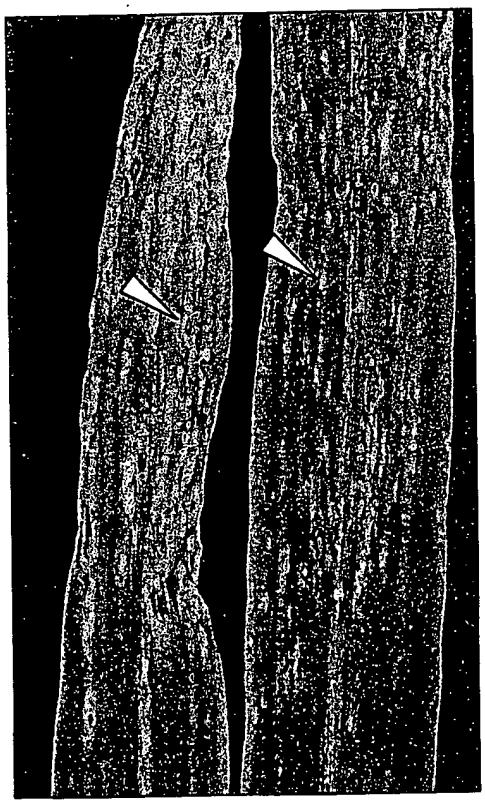


FIG. 12

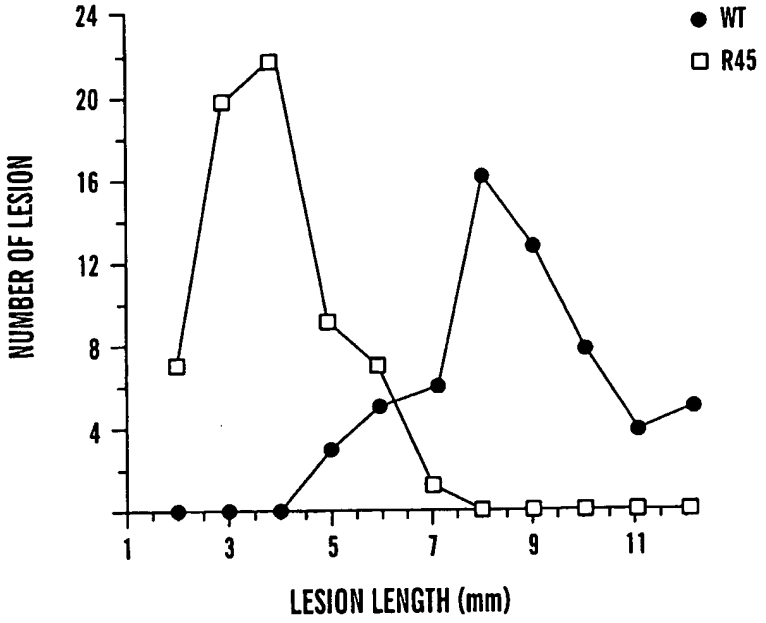


FIG. 13A

STRAIN	LESION SIZE (mm) FREQUENCY 1-4	5-8	9-12	MEAN	SD	
WT	0	52	48	8.5	1.0	A*
R45	86	14	0	3.5	0.9	B

* SIGNIFICANT DIFFERENCE AT $P<0.01$.

FIG. 13B

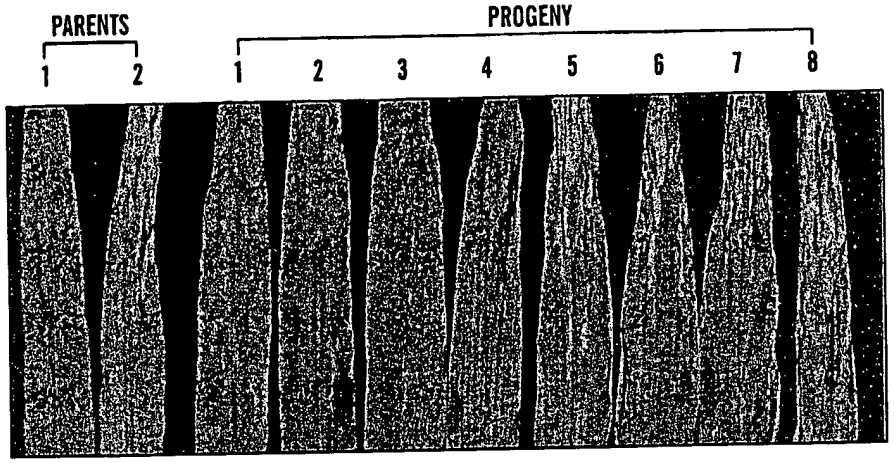


FIG. 14A

CROSS	PROGENY	PARENTAL TYPE path hygB^R	PARENTAL TYPE PATH hygB^S	NONPARENTAL TYPE path hygB^R	NONPARENTAL TYPE PATH hygB^S
R.C4.2696 X C5	RANDOM SPORES	24	22	0	0
1301-R33* X C5	tetrad1	4	4	0	0
	tetrad2	4	4	0	0
	tetrad3	4	4	0	0
	RANDOM SPORES	21	22	0	0

*13012-R33 (path, hygB^R, Tox, MAT-2) IS A PROGENY FROM THE FIRST CROSS, CARRYING THE R.C42696 MUTATION

FIG. 14B

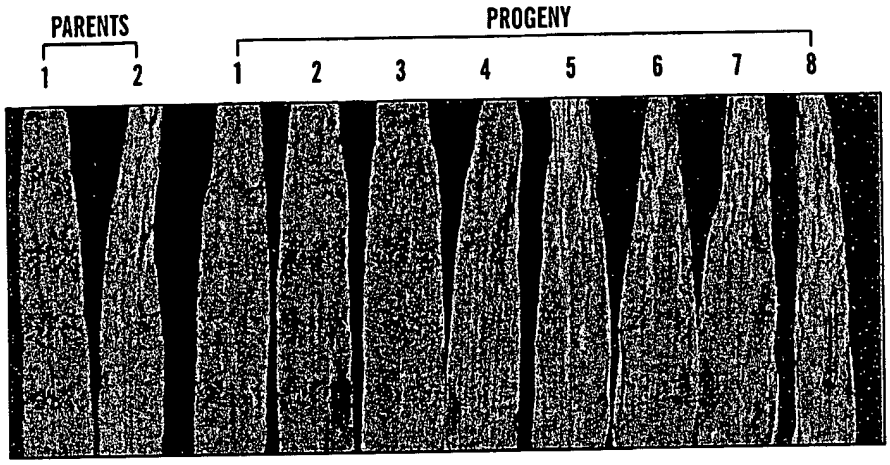


FIG. 14A

CROSS	PROGENY	PARENTAL TYPE path hygB^R	PARENTAL TYPE PATH hygB^S	NONPARENTAL TYPE path hygB^R	NONPARENTAL TYPE PATH hygB^S
R.C4.2696 X C5	RANDOM SPORES	24	22	0	0
1301-R33* X C5	tetrad1	4	4	0	0
	tetrad2	4	4	0	0
	tetrad3	4	4	0	0
	RANDOM SPORES	21	22	0	0

*13012-R33 (path, hygB^R, Tox, MAT-2) IS A PROGENY FROM THE FIRST CROSS, CARRYING THE R.C42696 MUTATION

FIG. 14B

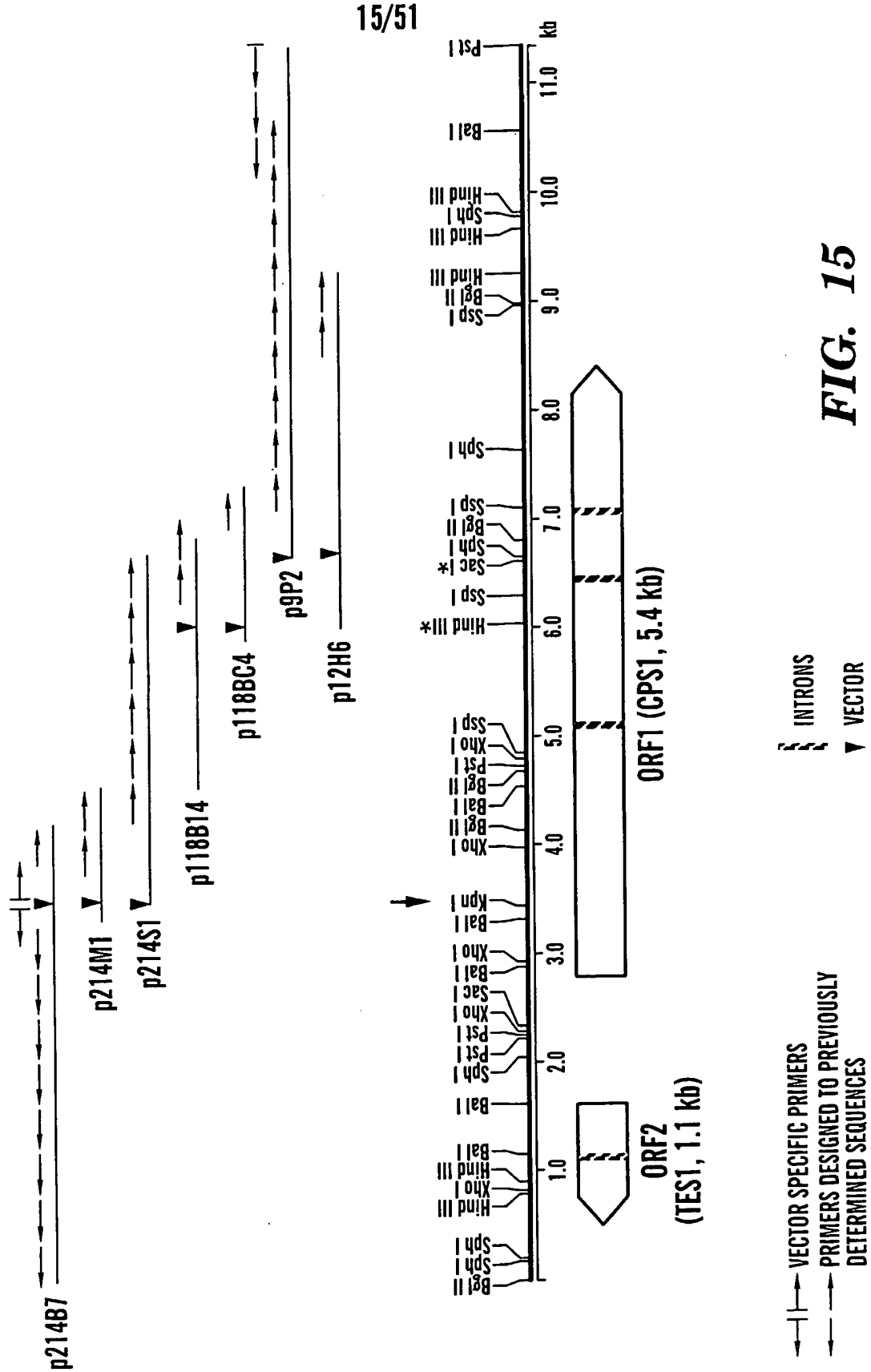

p214B7
p214M1
p214S1
p118B14
p118BC4
p9P2
p12H6
Pst I
Bal I
Hind III
Sph I
Hind III
Hind III
Bgl II
Ssp I
Sph I
Ssp I
Sac I*
Sph I
Bgl II
Ssp I
Hind III*
Ssp I
Xho I
Pst I
Bgl II
Bal I
Bgl II
Xho I
Kpn I
Bal I
Xho I
Bal I
Sac I
Xho I
Pst I
Pst I
Sph I
Bal I
Hind III
Xho I
Bal I
Hind III
Bgl II
Sph I
Sph I
kb
11.0
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ORF1 (CPS1, 5.4 kb)
ORF2
(TES1, 1.1 kb)
INTRONS
VECTOR
VECTOR SPECIFIC PRIMERS
PRIMERS DESIGNED TO PREVIOUSLY
DETERMINED SEQUENCES

FIG. 15

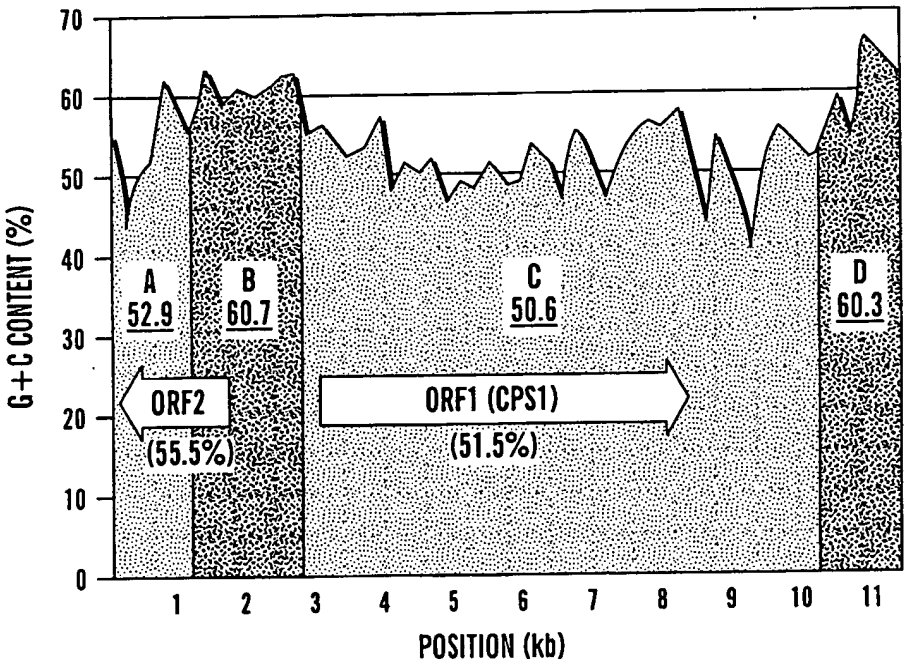

70
60
50
40
30
20
10
0
G+C CONTENT (%)
A
52.9
B
60.7
C
50.6
D
60.3
ORF2
(55.5%)
ORF1 (CPS1)
(51.5%)
1
2
3
4
5
6
7
8
9
10
11
POSITION (kb)

FIG. 16

```
TGCCTGCGCCTGTGCTTGTGCCTGTGGAATGTCGCGGCCCGCTGCTGCAT      -776

AGCCTATCTGTACATACAACACCATCCCATCCCGCTTCACCTGCCTTGCC      -726

TCCCTCCTCGTGCCACACATCCGCCGCCCACAACACCATGGCTGCGACCA      -676

ACCCCGAGCTGCAGGCCAAACTGCAGGAGCTGGACCACGAGCTCGAGGAG     -626

GGCGATATTACACAAAAAGGGTCCGTACTGCTGCACCACCACCGCCATCC      -576

GCCTCTCTGCGTGCGCTAATCAGTCGCATAGCTATGAAAAACGTCGCACC      -526

GTGCTGCTGTCGCAGTATCTAGGGCCTGACTTTGCTGCCCAGTTGCAGGC      -476

CGACCTGAACCAGCAGAACCCACCCCAACCATCCAGTGAGGGCTCTCGCT      -426

CCCGCACCGCATCCTTTGCTATTCCGTCCGGTCCGAGTCCATCACGGCGA      -376

CCACAACCCCCACATATCCAGCTCCCCCGCCCCGACTCATACCATGACGC      -326

TTCCGCACAGGGCCAATTGGGCGCACCCATGCCATATGCGAACGCCTCCG      -276

CCGCTGCCTCGGGGGGCTCGCAGTACATGGCATACCCGCCCAGCCAAGTC      -226

GGCCGTTTTCAAGAGAAGCAGCTGGGCCTGCGTACAAATTCGCTCCAGCG      -176

CAATTCCTCACAGCTGTCGCAAGGAAGCGAGACGTTCATTCCACGGCCTC      -126

AAACGCCTGAATACAACCACTCGCGCGAGCCCACCATGATGGGCAACTAC       -76

GCCTTCAATCCAGACAATCAGCAAAGTTATGATGGCCAATTTGGCTCTCC       -26

GGGAGAGGCCAGTCGAAGGAGCACCATGCTCGAGGTAAACCAGGGTTATT        25
                              M  L  E  V  N  Q  G  Y

TTTCCGACTTCACAGGCCAGCAGATGCAAGACAATCGCGACTCGTATGGG        75
F  S  D  F  T  G  Q  Q  M  Q  D  N  R  D  S  Y  G

GGACCCAACCGCTACTCGTCGGGAGATGCCTTTTCTCCTACCGCCGCGAT       125
 G  P  N  R  Y  S  S  G  D  A  F  S  P  T  A  A  I

TCCACCTCCCATGATGAACCCCAACGATCTCCCCTTGGGCGCTGCTGAAA       175
  P  P  P  M  M  N  P  N  D  L  P  L  G  A  A  E

CCATGATGCCGCTAGAGCCCCGCGATCTGCCTTTTGACGTTTACGACCCT       225
T  M  M  P  L  E  P  R  D  L  P  F  D  V  Y  D  P
```

FIG. 17A

```
CACAACCCCAATGTCAAAATGTCAAAGTTTGACAACATTGGCGCTGTCTT  275
 H  N  P  N  V  K  M  S  K  F  D  N  I  G  A  V  L

GCGTCACCGAAGTCGCACACAGCCAAGGACGACTGCCTTCTGGGTCCTTG  325
   R  H  R  S  R  T  Q  P  R  T  T  A  F  W  V  L

ACGCAAAAGGCAAAGAGACGGCGTCCATCACCTGGGAAAAGGTGGCTAGT  375
 D  A  K  G  K  E  T  A  S  I  T  W  E  K  V  A  S

CGCGCGGAAAAGGTGGCCAAAGTGATTCGGGACAAGAGCAACCTCTATCG  425
   R  A  E  K  V  A  K  V  I  R  D  K  S  N  L  Y  R

AGGCGACCGTGTGGCATTAGTGTACAGGGATACAGAAATCATTGATTTTG  475
   G  D  R  V  A  L  V  Y  R  D  T  E  I  I  D  F

TCGTGGCGTTGATGGGCTGCTTCATTGCGGGCGTTGTAGCGGTACCCATC  525
 V  V  A  L  M  G  C  F  I  A  G  V  V  A  V  P  I
                   Core 1

AATAGCGTCGACGACTACCAGAAACTCATTCTTCTCCTAACGACAACTCA  575
 N  S  V  D  D  Y  Q  K  L  I  L  L  L  T  T  T  Q

AGCTCATCTCGCATTGACCACAGACAACAATCTCAAGGCCTTTCATCGTG  625
   A  H  L  A  L  T  T  D  N  N  L  K  A  F  H  R

ACATTAGTCAGAACCGTCTGAAATGGCCGAGTGGGGTAGAGTGGTGGAAG  675
 D  I  S  Q  N  R  L  K  W  P  S  G  V  E  W  W  K

ACGAACGAGTTTGGCAGCCACCACCCCAAGAAACATGACGATACTCCAGC  725
   T  N  E  F  G  S  H  H  P  K  K  H  D  D  T  P  A

TTTGCAAGTACCAGAGGTTGCCTATATTGAGTTCTCGCGTGCACCTACTG  775
   L  Q  V  P  E  V  A  Y  I  E  F  S  R  A  P  T
                                 Core 2

GTGACCTTCGCGGTGTGGTGCTTAGTCACCGGACTATTATGCACCAAATG  825
 G  D  L  R  G  V  V  L  S  H  R  T  I  M  H  Q  M

GCCTGCATCAGTGCCATGATTAGCACGATACCCACCAACGCTCAGAGCCA  875
   A  C  I  S  A  M  I  S  T  I  P  T  N  A  Q  S  Q

AGACACGTTCAGCACTAGCCTACGGGATGCAGAGGGAAAGTTCGTTGCTC  925
   D  T  F  S  T  S  L  R  D  A  E  G  K  F  V  A

CAGCACCGTCCAGAAACCCCACAGAAGTGATCCTCACGTACCTCGACCCG  975
 P  A  P  S  R  N  P  T  E  V  I  L  T  Y  L  D  P

CGCGAAAGCGCTGGTCTCATTCTCAGTGTCTTGTTTGCAGTTTATGGAGG 1025
   R  E  S  A  G  L  I  L  S  V  L  F  A  V  Y  G  G
```

FIG. 17B

```
CCACACCACCGTATGGCTCGAGACAGCGACCATGGAAACCCCGGGTCTAT   1075
  H  T  T  V  W  L  E  T  A  T  M  E  T  P  G  L

ATGCACATCTCATCACCAAATACAAGTCCAACATACTGCTAGCGGATTAC   1125
Y  A  H  L  I  T  K  Y  K  S  N  I  L  L  A  D  Y

CCAGGCCTCAAGCGCGCTGCATACAACTACCAACAGGATCCAATGGCTAC   1175
P  G  L  K  R  A  A  Y  N  Y  Q  Q  D  P  M  A  T

AAGAAACTTCAAGAAAAACACAGAACCCAACTTCGCCTCCGTGAAGATCT   1225
  R  N  F  K  K  N  T  E  P  N  F  A  S  V  K  I

GTCTGATTGACACGCTTACCGTCGACTGTGAATTTCACGAAATTCTCGGA   1275
C  L  I  D  T  L  T  V  D  C  E  F  H  E  I  L  G

GATCGATATTTCAGGCCACTGCGAAACCCTAGAGCGCGAGAACTGATCGC   1325
 D  R  Y  F  R  P  L  R  N  P  R  A  R  E  L  I  A

GCCAATGCTCTGCTTGCCAGAACATGGTGGAATGATAATATCTGTACGCG   1375
  P  M  L  C  L  P  E  H  G  G  M  I  I  S  V  R

ACTGGCTAGGTGGAGAGGAGCGCATGGGCTGCCCGCTAAGCATAGCAGTA   1425
D  W  L  G  G  E  E  R  M  G  C  P  L  S  I  A  V

GAAGAGTCAGATAATGATGAAGATGATACAGAGGATAAGTATGCAGCGGC   1475
 E  E  S  D  N  D  E  D  D  T  E  D  K  Y  A  A  A

AAATGGCTACTCCAGTCTTATTGGTGGTGGCACTACAAAGAACAAAAAGG   1525
  N  G  Y  S  S  L  I  G  G  G  T  T  K  N  K  K

AGAAGAAGAAGAAAGGCCCGACAGAGCTTACAGAAATCTTGCTGGACAAG   1575
E  K  K  K  K  G  P  T  E  L  T  E  I  L  L  D  K

GAAGCTCTGAAGATGAACGAAGTCATTGTTCTGGCCATTGGAGAAGAAGC   1625
 E  A  L  K  M  N  E  V  I  V  L  A  I  G  E  E  A

AAGCAAGCGGGCAAACGAGCCCGGCACCATGCGAGTCGGTGCCTTTGGAT   1675
  S  K  R  A  N  E  P  G  T  M  R  V  G  A  F  G

ACCCCATACCGGATGCGACACTAGCTATTGTAGACCCTGAGACAAGTCTT   1725
Y  P  I  P  D  A  T  L  A  I  V  D  P  E  T  S  L

CTATGTTCACCATACTCGATAGGCGAGATCTGGGTAGATTCGCCTTCACT   1775
  L  C  S  P  Y  S  I  G  E  I  W  V  D  S  P  S  L
                       Core 3
CTCTGGTGGCTTCTGGCAGCTGCAGAAGCATACAGAGACCATTTTCCATG   1825
  S  G  G  F  W  Q  L  Q  K  H  T  E  T  I  F  H
```

FIG. 17C

CTCGACCATACCGTTTCGTTGAGGGTAGCCCTACGCCACAGTTGCTTGAA 1875
A R P Y R F V G G S P T P Q L L E

CTCGAGTTTCTGCGTACTGGACTCCTCGGCTTTGTTGTAGAGGGAAAAAT 1925
L E F L R T G L G F V V E G K I
Core 4

ATTTGTCCTTGGACTGTACGAAGATCGCATCAGACAGCGTGTTGAATGGG 1975
F V L G L Y E D R I R Q R V E W
Core 5

TAGAAAATGGTCAGCTTGAAGCCGAGCATCGATACTTTTTTGTGCAGCAC 2025
Y E N G Q L E A E H R Y F F V Q H

CTGGTCACAAGCATTATGAAGGCCGTGCCAAAAATTTACGACTGgtaagt 2075
L V T S I M K A V P K I Y D C

gagctgccaacagagcaaggactgtctaacgtgtcatagCTCGTCGTTTG 2125
S S F

ATTCTTATGTAAATGGTGAATACCTGCCAATCATTCTCATCGAGACGCAG 2175
D S Y V N G E Y L P I I L I E T Q

GCCGCATCGACTGCGCCCACAAACCCAGGTGGACCACCACAACAATTGGA 2225
A A S T A P T N P G G P P Q Q L D

TATACCATTTTTGGATTCACTATCTGAGAGGTGCATGGAGGTCCTTTACC 2275
I P F L D S L S E R C M E V L Y
Core 6

AAGAGCATCATTTACGGGTATACTGCGTGATGATTACAGCACCTAATACA 2325
Q E H H L R V Y C V M I T A P N T

CTTCCACGAGTCATCAAGAACGGACGGCGAGAAATTGGCAATATGCTGTG 2375
L P R V I K N G R R E I G N M L C

TAGGAGAGAGTTTGACAATGGCTCTCTGCCCTGTGTNCACGTNAAGTTTG 2425
R R E F D N G S L P C V H V K F

GCATTGAGCGATCAGTGCAGAACATTGCGCTCGGTGACGATCCCGCTGGC 2475
G I E R S V Q N I A L G D D P A G

GGCATGTGGTCATTTGAGGCATCAATGGCACGTCAGCAATTCTTGATGCT 2525
G M W S F E A S M A R Q Q F L M L

CCAAGACAAGCAATACTCTGGTGTCGATCATCGCGAAGTCGTCATTGACG 2575
Q D K Q Y S G V D H R E V V I D

ACAGGACATCGACTCCACTCAATCAGTTCTCGAATATCCACGACCTGATG 2625
D R T S T P L N Q F S N I H D L M

FIG. 17D

```
CAATGGCGTGTATCTCGGCAGGCCGAGGAACTTGCTTACTGCACTGTCGA   2675
 Q  W  R  V  S  R  Q  A  E  E  L  A  Y  C  T  V  D

CGGTCGAGGAAAAGAGGGCAAAGGCGTCAATTGGAAGAAGTTTGATCAAA  2725
   G  R  G  K  E  G  K  G  V  N  W  K  K  F  D  Q

AGGTTGCGGGCGTAGCAATGTACCTCAAGAACAAGGTCAAGGTCCAGGCC  2775
K  V  A  G  V  A  M  Y  L  K  N  K  V  K  V  Q  A

GGCGATCATCTCCTTCTGATGTACACGCATTCAGAAGAATTTGTTTATGC   2825
   G  D  H  L  L  L  M  Y  T  H  S  E  E  F  V  Y  A

TGTTCATGCATGTTTTGTGCTTGGAGCTGTTTGCATACCAATGGCGCCAA   2875
   V  H  A  C  F  Y  L  G  A  V  C  T  P  M  A  P
                Core 1
TTGATCAGAACCGGTTGAATGAGGATGCGCCGGCCTTGCTGCATATCCTT   2925
I  D  Q  N  R  L  N  E  D  A  P  A  L  L  H  I  L

GCAGATTTCAAGGTCAAAGCCATTCTTGTCAACGCTGACGTTGACCATCT   2975
   A  D  F  K  V  K  A  I  L  V  N  A  D  V  D  H  L

GATGAAGATCAAGCAAGTATCGCAGCACATCAAACAATCGGCCGCTATCC   3025
   M  K  I  K  Q  V  S  Q  H  I  K  Q  S  A  A  I

TCAAGATCAGTGTGCCAAACACATACAGCACAACAAAGCCGCCAAAGCAA   3075
L  K  I  S  V  P  N  T  Y  S  T  T  K  P  P  K  Q

TCCAGTGGCTGCCGCGACCTCAAGCTTACAATTCGACCGGCATGGATTCA   3125
   S  S  G  C  R  D  L  K  L  T  I  R  P  A  W  I  Q

GGCGGGTTTCCCAGTGCTAGTCTGGACATACTGGACGCCCGATCAACGTC   3175
   A  G  F  P  V  L  V  W  T  Y  W  T  P  D  Q  R
                Core 2
GTATCGCAGTTCAGCTGGGCCATAGCCAAATCATGGCACTGTGCAAGGTC   3225
R  I  A  V  Q  L  G  H  S  Q  I  M  A  L  C  K  V

CAAAAAGAAACATGCCAAATGACAAGTACACGACCAGTCCTTGGTTGTGT   3275
   Q  K  E  T  C  Q  M  T  S  T  R  P  V  L  G  C  V

CCGGAGCACGATAGGACTTGGTTTCCTTCACACTTGTCTCATGGGAATCT   3325
   R  S  T  I  G  L  G  F  L  H  T  C  L  M  G  I

TCCTTGCCGCACCCACATACCTGGTGTCACCTGTTGACTTTGCACAAAAC   3375
F  L  A  A  P  T  Y  L  V  S  P  V  D  F  A  Q  N

CCTAATATTCTGTTCCAAACGCTTTCGCGGTACAAGATCAAGGATGCATA   3425
   P  N  I  L  F  Q  T  L  S  R  Y  K  I  K  D  A  Y
```

FIG. 17E

TGCAACGAGTCAAATGTTGGACCACGCCATCGCACGCGGAGCTGGTAAGA 3475
A T S Q M L D H A I A R G A G K

GTATGGCTCTGCACGAGCTGAAGAATCTCATGATTGCGACTGATGGAAGA 3525
S M A L H E L K N L M I A T D G R

CCACGCGTTGATGTTTgtaagtgaacatttgtatgagaggactttcatga 3575
P R V D V

ttgctaactcaatgcagACCAAAGAGTGCGTGTGCACTTTGCGCCAGCCA 3625
Y Q R V R V H F A P A

ACTTAGACCCAACCGCAATCAACACTGTCTACTCACATGTATTGAACCCA 3675
N L D P T A I N T V Y S H V L N P

ATGGTAGCATCACGATCATACATGTGTATTGAGCCAGTCGAGCTCCATCT 3725
M V A S R S Y M C I E P V E L H L

CGATGTGCATGCTCTGCGACGCGGCCTCGTCATGCCCGTTGACCCTGACA 3775
D V H A L R R G L V M P V D P D

CAGAGCCCAACGCTTTGCTCGTCCAAGACTCGGGCATGGTGCCAGTGAGC 3825
T E P N A L L V Q D S G M V P V S

ACGCAAATATCCATTGTCAACCCAGAGACCAACCAACTGTGCTTGAACGG 3875
T Q I S I V N P E T N Q L C L N G

CGAGTACGGCGAGATCTGGGTGCAGTCCGAGGCGAATGCTTATAGCTTCT 3925
E Y G E I W V Q S E A N A Y S F
Core 3

ACATGTCGAAAGAGCGCTTGGATGCAGAACGCTTCAATGGGAGGACGATT 3975
Y M S K E R L D A E R F N G R T I

GACGGAGACCCAAATGTGCGATATGTTCGTACAGGCGATTTAGGATTTTT 4025
D G D P N V R Y V R I G D L G F L
Core 4

GCACAGCGTGACACGGCCCATTGGACCCAACGGTGCACCTGTTGATATGC 4075
H S V T R P I G P N G A P V D M

AGGTGCTTTTCGTGCTTGGAAGCATAGGTGACACTTTTGAAGTCAACGGA 4125
Q V L F V L G S I G D T F E V N G

CTGAACCATTTCTCTATGGACATTGAGCAGTCTGTTGAACGTTGTCACCG 4175
L N H F S M D I E Q S V E R C H R

FIG. 17F

```
GAATATTGTCCCTGGAGGCTGgtacgtttcttcgattcgctgttatttag          4225
  N  I  V  P  G  G  C

taaatacttactaacactctacagTGCTGTTTTCCAGGCAGGTGGGCTTG          4275
                          A  V  F  Q  A  G  G  L

TTGTTGTCGTTGTGGAAATCTTCCGACGCAACTTCCTCGCAAGCATGGTG          4325
V  V  V  V  V  E  I  F  R  R  N  F  L  A  S  M  V

CCTGTGATTGTCAATGCAATTTTGAACGAGCATCAGCTGGTCATTGACAT          4375
 P  V  I  V  N  A  I  L  N  E  H  Q  L  V  I  D  I

TGTCTCGTTTGTGCAAAAGGGCGACTTCCACCGGTCTCGTCTGGGCGAGA          4425
   V  S  F  V  Q  K  G  D  F  H  R  S  R  L  G  E

AGCAACGCGGAAAGATTCTTGCAGGATGGGTCACACGGAAGATGCGCACA          4475
K  Q  R  G  K  I  L  A  G  W  V  T  R  K  M  R  T

ATAGCCCAGTACAGTATACGGGATCCTAATGGACAGGATTCCCAGATGAT          4525
 I  A  Q  Y  S  I  R  D  P  N  G  Q  D  S  Q  M  I
                              Core 6

CACGGAAGAGCCTGGTCCACGGGCTAGCATGACTGGAAGTATGCTTGGGC          4575
   T  E  E  P  G  P  R  A  S  M  T  G  S  M  L  G

GAATGGGCGGCCCAGCCAGTATCAAGGCCGGGTCGACAAGAGCACCGAGT          4625
R  M  G  G  P  A  S  I  K  A  G  S  T  R  A  P  S

CTAATGGGCATGACAGCGACTATGAATAATCTATCCCTTACACAGCAGCA          4675
 L  M  G  M  T  A  T  M  N  N  L  S  L  T  Q  Q  Q

ACAGCAGCAATACCAACAGCCGGGTATGTATGCTCAACAGCAAGGCATGC          4725
   Q  Q  Q  Y  Q  Q  P  G  M  Y  A  Q  Q  Q  G  M

ACCCCCAGCAACAACACCAATTTAGCATGTCCAACACGCCACCACAAGGT          4775
H  P  Q  Q  Q  H  Q  F  S  M  S  N  T  P  P  Q  G

CCACCCCAAGGCGTAGAACTACATGATCCTAGCGACCGCACACCAACAGA          4825
 P  P  Q  G  V  E  L  H  D  P  S  D  R  T  P  T  D

CAACCGGCACTCTTTCCTTGCCGACCCGCGTATGCAGAACCAGGGCCAAA          4875
   N  R  H  S  F  L  A  D  P  R  M  Q  N  Q  G  Q

TGAACGAGACGGGCGCCTACGAACCCATGAACTATCAAAACGCGTATCAT          4925
M  N  E  T  G  A  Y  E  P  M  N  Y  Q  N  A  Y  H

CCGCATCAACAACAATACGAATCTGAAGACGGGGGGAGCAGACTCAGCGG          4975
 P  H  Q  Q  Q  Y  E  S  E  D  G  G  S  R  L  S  G
```

FIG. 17G

```
CCCCGTGCCAGACGTGCTGCGGCCGGGTCCTTCATCCGGGTCCATAGAGC      5025
  P  V  P  D  V  L  R  P [G  P  S  S  G] S  I  E

AGCACGACCAAGCTAACAACGACAACAATATGTGGAATAATCGCGAGTAC      5075
Q  H  D  Q  A  N  N  D  N  N  M  W  N  N  R  E  Y

TATGGTAACAGCCCATCGTATGCAGGCGGATACACGCAAGATGGCAATAT      5125
 Y  G  N  S  P  S  Y  A  G  G  Y  T  Q  D  G  N  I

CCACGAGCAGCAACAACACGATGAGTACACGAGTAATGCGTCATATGGCG      5175
  H  E  Q  Q  Q  H  D  E  Y  T  S  N  A  S  Y  G

GAAATCAAGGAGCAGGCGGAGGCAGCGGCGGCGGTGGCGGTCTCCGAGTT      5225
G  N  Q  G  A  G  G  G  S  G  G  G  G  G  L  R  V

GCAAATCGTGACAGCTCCGACAGCGAGGGTGCAGATGACGACGCTTGGAG      5275
 A  N  R  D  S  S  D  S  E  G  A  D  D  D  A  W  R

ACGTGATGCCCTTGCTCAGATCAATTTTGCGGGCGGCGCTGCTGCTGCCT      5325
  R  D  A  L  A  Q  I  N  F  A  G  G  A  A  A  A

CCGCTGGAGCACCTGCTGCTGGTGCTTCTTCTTCGCAGCCGGGCCATGCG      5375
S  A  G  A  P  A  A  G  A  S  S  S  Q  P  G  H  A

CAGTAGACGGGATATGCGTGAGTTTTTTTTTAAATTTCGTACATAGAGAC      5425
 Q END

CGTTGTATACGCAGGTTTCAAATTAGAAGAGCGAATATGCATATCAGCTG      5475

TTGTTCAATGTTCTAGTTTGGGAAGGTTAACCCCCCCCCCTTCCCCTTCC      5525

AAGACTTTTCACTTGTTTGTGTGTGATTTAAATCTGGAGATTTCAAATCT      5575

ACATCTCGCTATACATAGGTGTTGTTTGATAACGTAGGGGGCAGAAGGGT      5625

ATCTCGTGATATTAGACTGGGAGTTGCATGAATCAAGGTGTTGAGCAAAA      5675

AAAGAGAGAGCGGTGAAGGGCGGGGGGGATAGGTGGTGTGCACGTGGCTG      5725
```

FIG. 17H

```
AAGCCAGGTACGCATCTGCGCAGGGCCCGGCAAGCCCAGTATGACGTTTT    -286

TACCCCAGAGCCAGCCTGCTCCTCGGCTTGGCGCCGCCTGGCTTAGTCAG    -236

CCCCCATCAGTCAGCGGCCAGTCACGTGTTCGGCGCGACAAGCTCCCACA    -186

TGGCCGATGATCATTGTGCCTCGGCTCATGCGCGTCTGCGCTGCTTTGAC    -136

GTGCCCCGGAAATGACGACGCCCAAAGTCTCACACAGCCGCTCGCTGCTG     -86

                          **********
CTGCTTGCATCTGGTCTTGCTACACTGGCCTCCTTTTCGCAGAACACAGA     -36

ACGATCCACCAACAGGAAAACTATCAGCCACCACCATGGCCACCCTCATC      15
                                    M  A  T  L  I

CGCCCCCCGCCCGCCGACCACTCCCAGTCGGCCATTGAGAACGTCCTTGA      65
 R  P  P  P  A  D  H  S  Q  S  A  I  E  N  V  L  E

ACTGACGCAGCTGGCCGACATTGATCCCAACATCTTCACAAACACTCGTC     115
  L  T  Q  L  A  D  I  D  P  N  I  F  T  N  T  R

CCCTATGGCATCCCCCCGGTGCGCGCGGCATCTTTGGCGGTGCTGCCATT     165
P  L  W  H  P  P  G  A  R  G  I  F  G  G  A  A  I

GCCCAGACACTCAGCGCCGCCCAGAAGACGGTTGACCCCGACTTCACTGT     215
 A  Q  T  L  S  A  A  Q  K  T  V  D  P  D  F  T  Y

GCATTCGATGCACTGCTACTTCATCCTTGCAGGCAACTCCGAGATCCCCG     265
 H  S  M  H  C  Y  F  I  L  A  G  N  S  E  I  P

TCATATACCATGTAGAAAGGGTGCGCTCGGGAAAGTCGTTCGCAACCAGG     315
V  I  Y  H  V  E  R  V  R  S  G  K  S  F  A  T  R

ACGGTGCAGGCCCGGCAAAGGGGCAACGTCATCTTCACTACGACCATGAG     365
 T  V  Q  A  R  Q  R  G  N  V  I  F  T  T  T  M  S

TTTTGTTCGCCAGAACAGTGGAGGCGCTCAAAAGGTCGAGCACATCTACC     415
  F  V  R  Q  N  S  G  G  A  Q  K  V  E  H  I  Y

CGATGCCAGACGTCCCGGCGCCCAAGGAAGGTAGCGACGACTTGAAGACG     465
P  M  P  D  V  P  A  P  K  E  G  S  D  D  L  K  T

CCCAACGATGGCCAGAGTCCCTTCCAGACCCAGCTGTTGCCCATTGAAAA     515
 P  N  D  G  Q  S  P  F  Q  T  Q  L  L  P  I  E  N

CGgtgagtgacttcgcataccatccttaagcctgccctagctcatatcta     565
```

FIG. 18A

```
gCAGACGACTCCGACAAGCCCCACACCAAGAAATGCCGACAATGGATAAA      615
 A  D  D  S  D  K  P  H  T  K  K  C  R  Q  W  I  K

GGCTCGCGGTAAAATCTCCCCCGCTGGCGGTCACGAAGCCCATCTCTCCG      665
   A  R  G  K  I  S  P  A  G  G  H  E  A  H  L  S

CCATCGCATACATGTCCGATAGCTACTTCATCGGCACCGTCGCGCGTGCG      715
A  I  A  Y  M  S  D  S  Y  F  I  G  T  V  A  R  A

CACAAGCTTCTGCGCTACTCGAACCAGCGCAAGAGCAGGGCCAGGTCGAG      765
 H  K  L  L  R  Y  S  N  Q  R  K  S  R  A  R  S  S

CATCGACGAGGACGTACTTAAGAAGCTGCTCGAGATGGATGATGCCGAGT      815
   I  D  E  D  V  L  K  K  L  L  E  M  D  D  A  E

TACAGCGCCAAAGCTTTGTCAACGAATCAGACAAGCAGCGCATACGTGAA      865
L  Q  R  Q  S  F  V  N  E  S  D  K  Q  R  I  R  E

TTGAGGAAAGCAGAAGACCTGGCAAAGAGCGGAGACGCAAAGCCTGAGAT      915
 L  R  K  A  E  D  L  A  K  S  G  D  A  K  P  E  I

TGGCATGATGGTTAGTCTGGACCACACCATCTACTTTCACAATCCTCGCA      965
   G  M  M  V  S  L  D  H  T  I  Y  F  H  N  P  R

GTTTCCGCGCAGATGAATGGATCTTCACCGAGATGGAGACCCCTTGGGCT     1015
S  F  R  A  D  E  W  I  F  T  E  M  E  T  P  W  A

GGTGATGGTCGCGGCCTCGTTTCCCAGAGAATGTATACCAAGGATGGCAC     1065
 G  D  G  R  G  L  V  S  Q  R  M  Y  T  K  D  G  T

GCTCATTGCCAGCTGCGTCCAAGAGGTAAGCAGCTTGCTTATGTCTGCTA     1115
   L  I  A  S  C  V  Q  E  V  S  S  L  L  M  S  A

TAGTAGACTGTGCTTACATGAATAGGGCGTCATACGATTGAAGCAGAACG     1165
I  V  D  C  A  Y  M  N  R  A  S  Y  D  END

AAAGTAAGCTATAGCCATGACCGTGTGTAGATTGCTCGTATGTCCGTGTA     1215

CTTTGATTCGATTGACTTTTATGCCGCCGCCTTTCACTTGACCTTGTCGG     1265

AGTAGACTCCGCTGTTTCTATTGTTTTAGTAGCTTTAGATTTGGAGGTCT     1315

GGAACACACGTACCTCACGTATACACTTTTAAATACAAGGTTCCTCGTTA     1365

CAAATAGCAACTCGCACTTGTGTAGTACACGACTTGGCACAGCAAGTGCA     1415

TGCAGATGATGCGCACTGCATGCAGCCCTAGAAGCCTGCACCTGTGCTGT     1465

TCGTCCTTCCCACGTTACAGGCCAGGGTCCAATTAGCGCAATGCCGGTCT     1515

CGGGCTAGCATGGCCCTCACACTTGGATTCTCCGAAGCTACCGTCGGCAG     1565
```

FIG. 18B

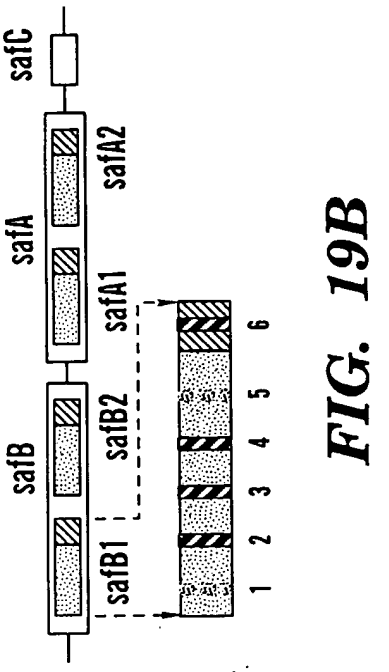


FIG. 19B

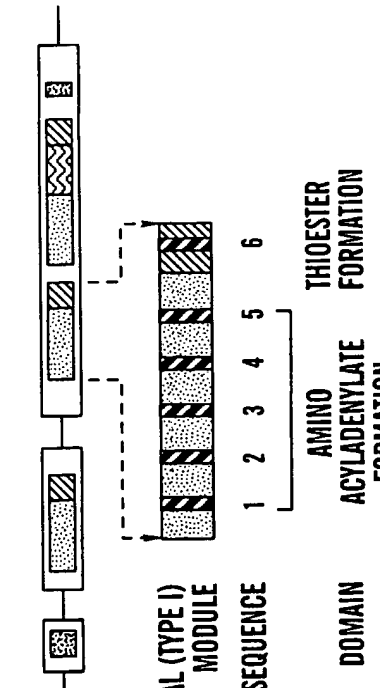


FIG. 19A

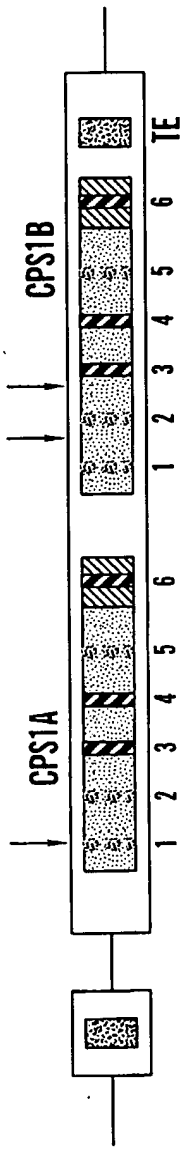


FIG. 19C

THIOESTERASE

N-METHYLTRANSFERASE (ABOUT 400 aa, IN TYPE II MODULE ONLY)

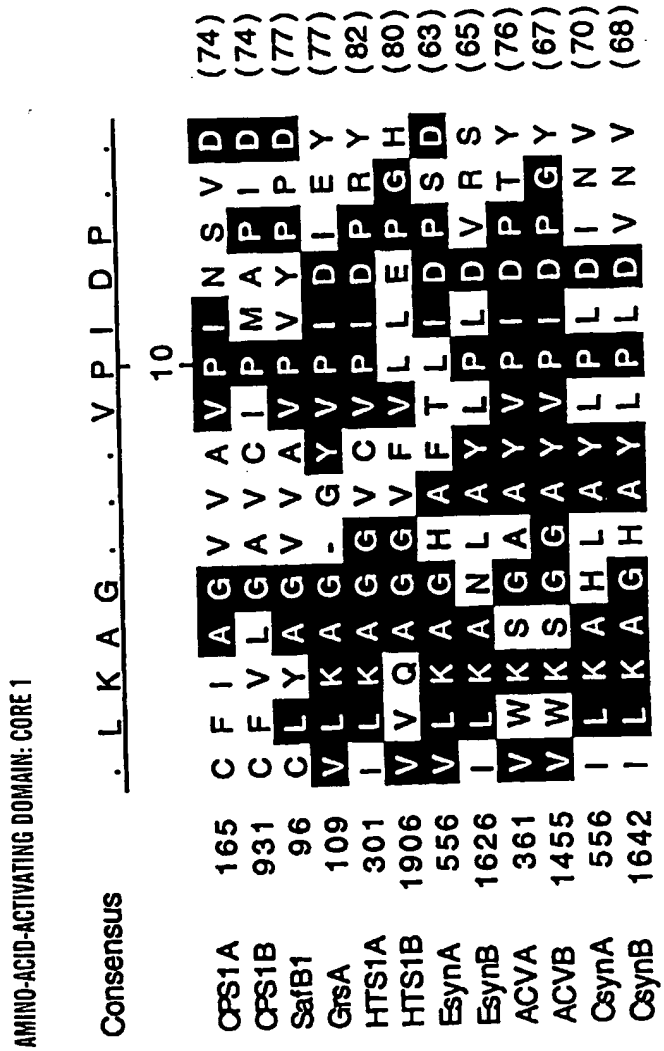
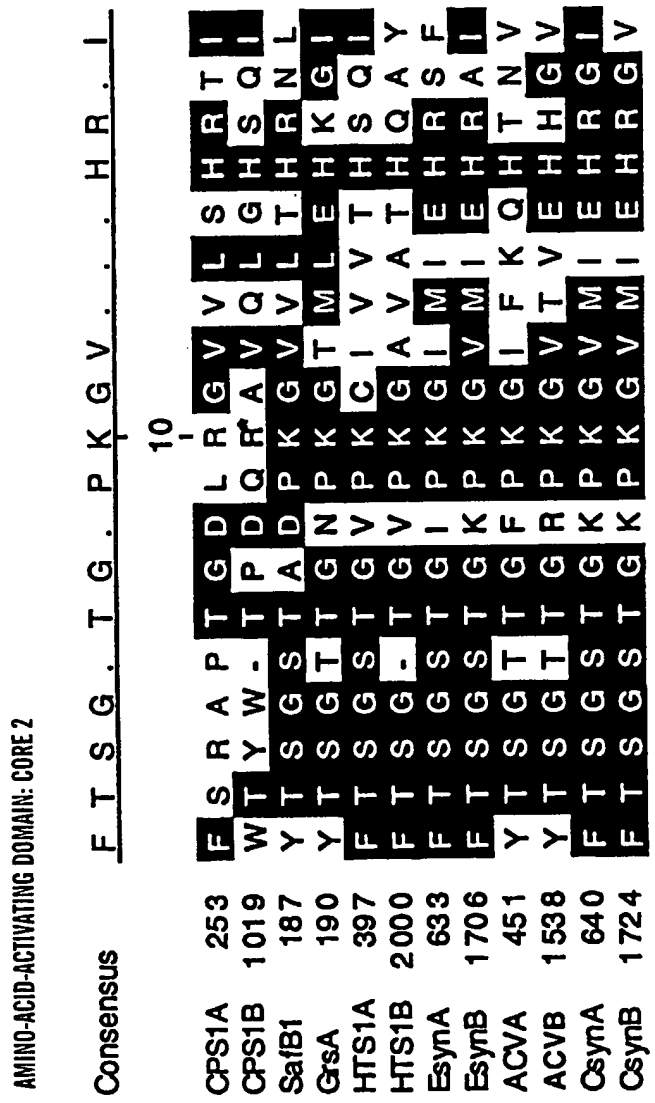

AMINO-ACID-ACTIVATING DOMAIN: CORE 1
Consensus
. L K A G . . . . V P I D P . .
10
CPS1A 165 C F I A G V V A V P I N S V D (74)
CPS1B 931 C F V L G A V C I P M A P I D (74)
SafB1 96 C L Y A G V A V P V Y P P D (77)
GrsA 109 V L K A G - G Y V P I D I E Y (77)
HTS1A 301 I L K A G G V C V P I D P R Y (82)
HTS1B 1906 V V Q A G G V F V L L E P G H (80)
EsynA 556 V L K A G H A F T L I D P S D (63)
EsynB 1626 I L K A N L A Y L P L D V R S (65)
ACVA 361 V W K S G A A Y V P I D P T Y (76)
ACVB 1455 V W K S G G A Y Y V P I D P G Y (67)
CsynA 556 I L K A H L A Y L P L D I N V (70)
CsynB 1642 I L K A G H A Y L P L D V N V (68)

FIG. 20A



* AN INSERTION (2 RESIDUES BETWEEN R AND A) IS NOT SHOWN.

FIG. 20B

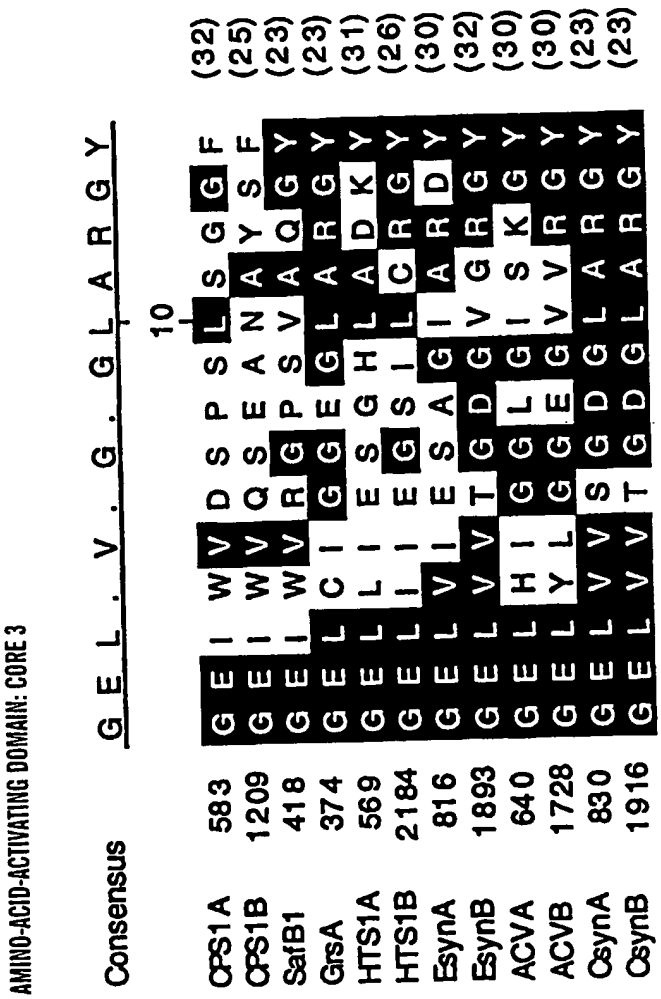

AMINO-ACID-ACTIVATING DOMAIN: CORE 3
Consensus G E L . V . G . G L A R G Y
10
CPS1A 583 G E I W V D S P S L S G G F (32)
CPS1B 1209 G E I W V Q S E A N A Y S F (25)
SafB1 418 G E I W V R G P S V A Q G Y (23)
GrsA 374 G E L C I G G E G L A R G Y (23)
HTS1A 569 G E L L I E S G H L A D K Y (31)
HTS1B 2184 G E L I I E G S I L C R G Y (26)
EsynA 816 G E L V I E S A G I A R D Y (30)
EsynB 1893 G E L V V T G D G V G R G Y (32)
ACVA 640 G E L H I G G L G I S K G Y (30)
ACVB 1728 G E L Y L G G E G V V R G Y (30)
CsynA 830 G E L V V S G D G L A R G Y (23)
CsynB 1916 G E L V V T G D G L A R G Y (23)

FIG. 20C

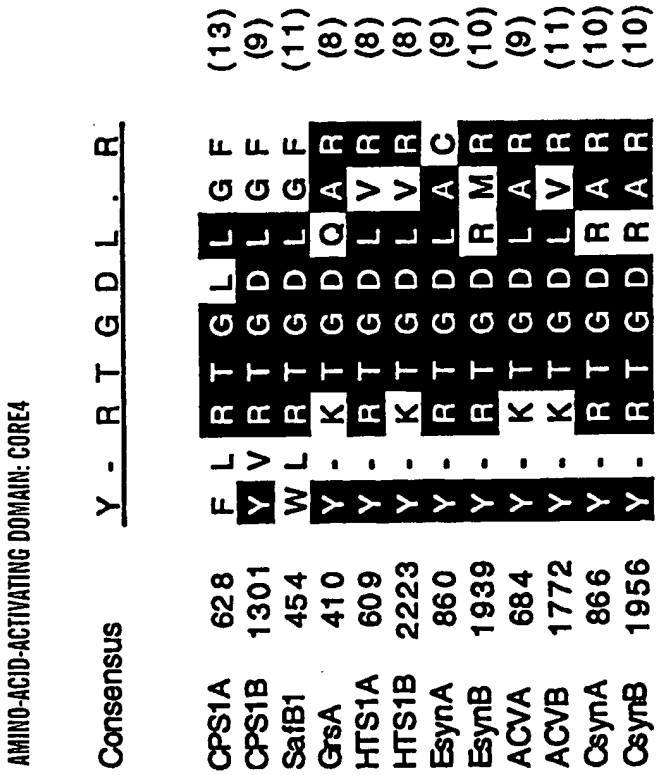
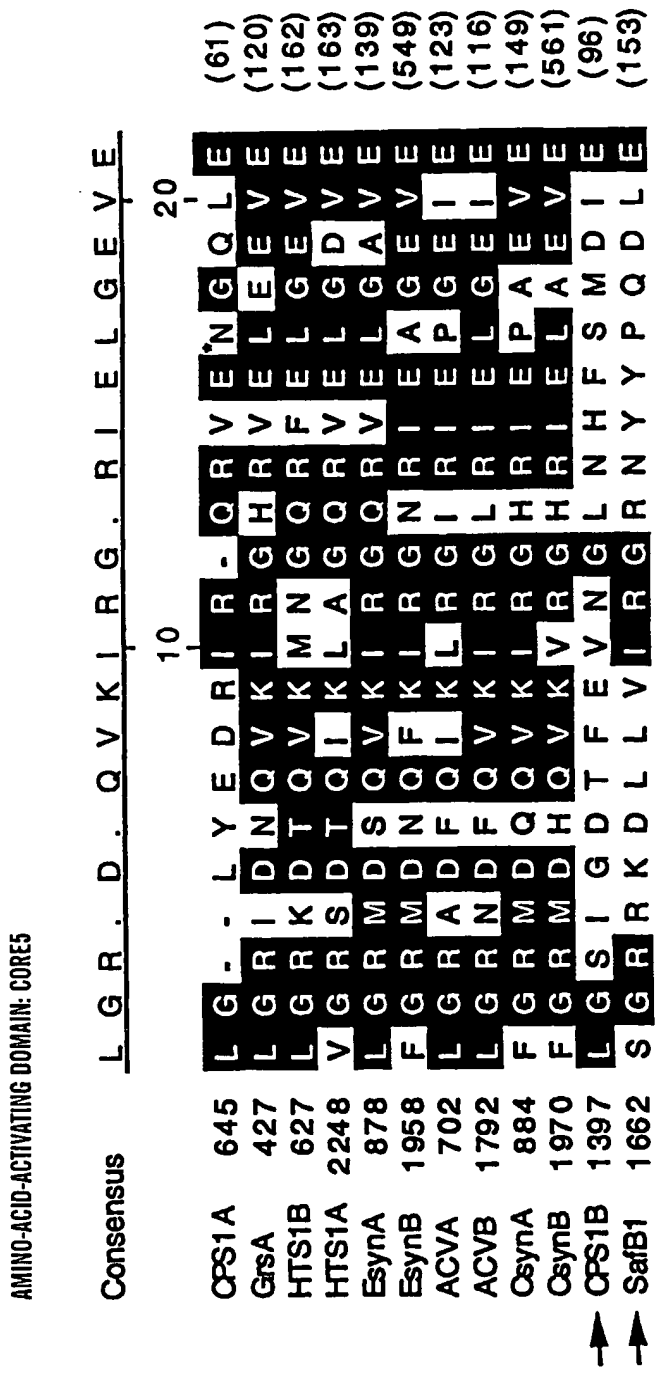
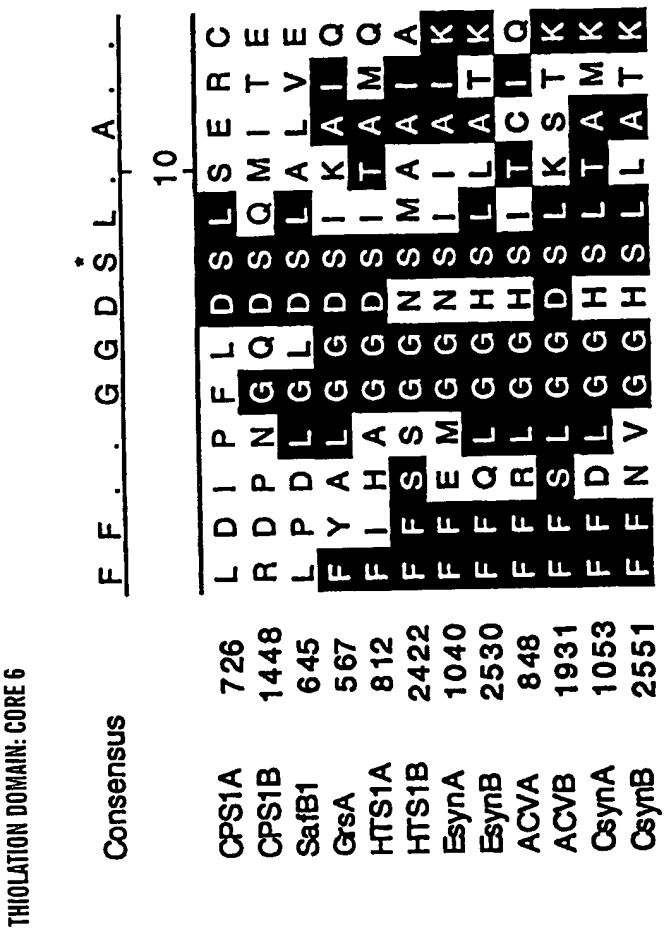

AMINO-ACID-ACTIVATING DOMAIN: CORE4
Consensus Y - R T G D L . R
CPS1A 628 F L R T G L L G F (13)
CPS1B 1301 Y V R T G D L G F (9)
SafB1 454 W L R T G D L G F (11)
GrsA 410 Y - K T G D Q A R (8)
HTS1A 609 Y - R T G D L V R (8)
HTS1B 2223 Y - K T G D L V R (8)
EsynA 860 Y - R T G D L A C (9)
EsynB 1939 Y - R T G D R M R (10)
ACVA 684 Y - K T G D L A R (9)
ACVB 1772 Y - K T G D L V R (11)
CsynA 866 Y - R T G D R A R (10)
CsynB 1956 Y - R T G D R A R (10)

FIG. 20D



*AN INSERTION (TWO AMINO ACID) BETWEEN E AND N IN CPS1A IS NOT SHOWN.
THE POORLY CONSERVED CORES 5 IN CPS1B AND SafB1 ARE INDICATED BY ARROWS

FIG. 20E



* ACTIVE SITE FOR 4'-PHOSPHOPANTETHEINE BINDING.
**TYPE II MODULES CONTAINING A METHYLTRANSFERASE DOMAIN (ABOUT 400 AMINO ACIDS) BETWEEN CORES 5 AND 6. ALL OTHERS ARE TYPE I MODULES WITHOUT THIS INSERTION.

FIG. 20F

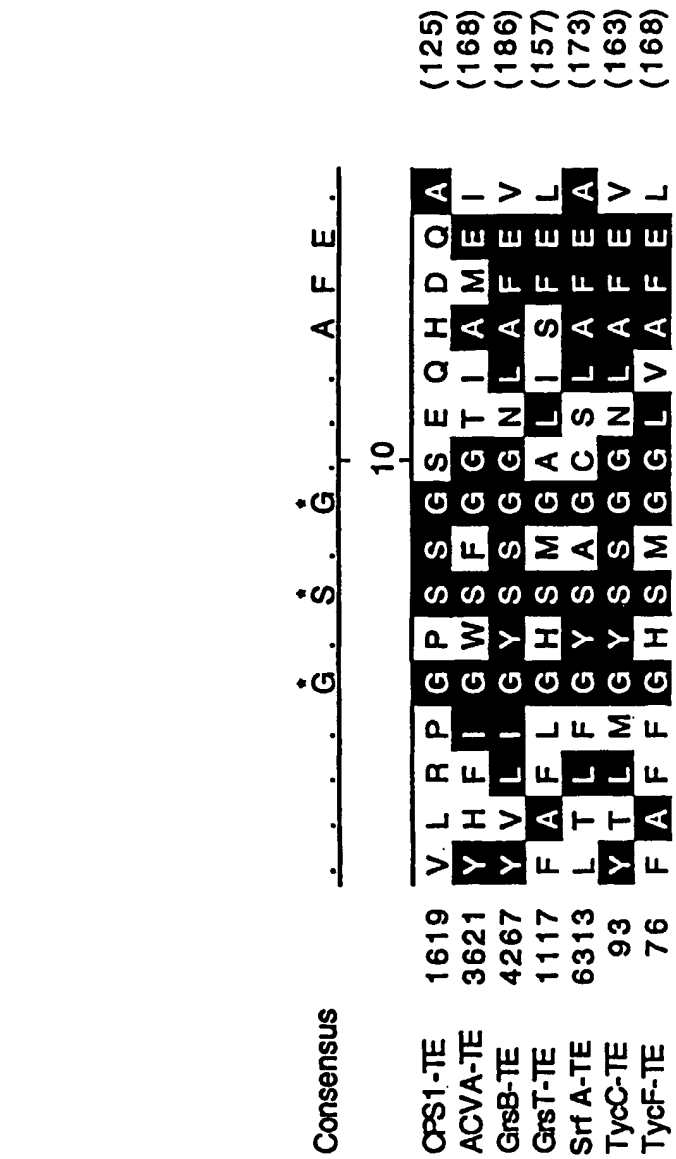

Consensus
CPS1-TE 1619 V L R P G P S S G S E Q H D Q A (125)
ACVA-TE 3621 Y H F I G W S F G G T I A M E I (168)
GrsB-TE 4267 Y V L I G Y S S G G N L A F E V (186)
GrsT-TE 1117 F A F L G H S M G A L I S F E L (157)
Srf A-TE 6313 L T L F G Y S A G C S L A F E A (173)
TycC-TE 93 Y T L M G Y S S G G N L A F E V (163)
TycF-TE 76 F A F F G H S M G G L V A F E L (168)
10

FIG. 21

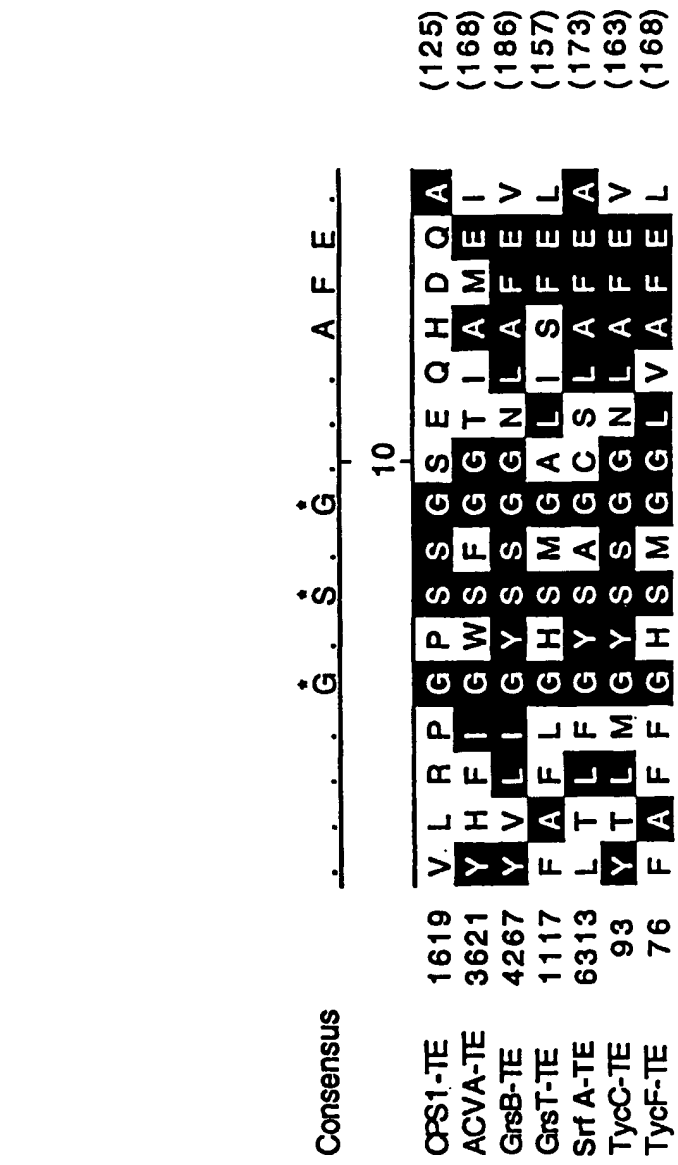

Consensus
CPS1-TE 1619 V L R P G P S S G S E Q H D Q A (125)
ACVA-TE 3621 Y H F I G W S F G G T I A M E I (168)
GrsB-TE 4267 Y V L I G Y S S G G N L A F E V (186)
GrsT-TE 1117 F A F L G H S M G A L I S F E L (157)
Srf A-TE 6313 L T L F G Y S A G C S L A F E A (173)
TycC-TE 93 Y T L M G Y S S G G N L A F E V (163)
TycF-TE 76 F A F F G H S M G G L V A F E L (168)
10

FIG. 21

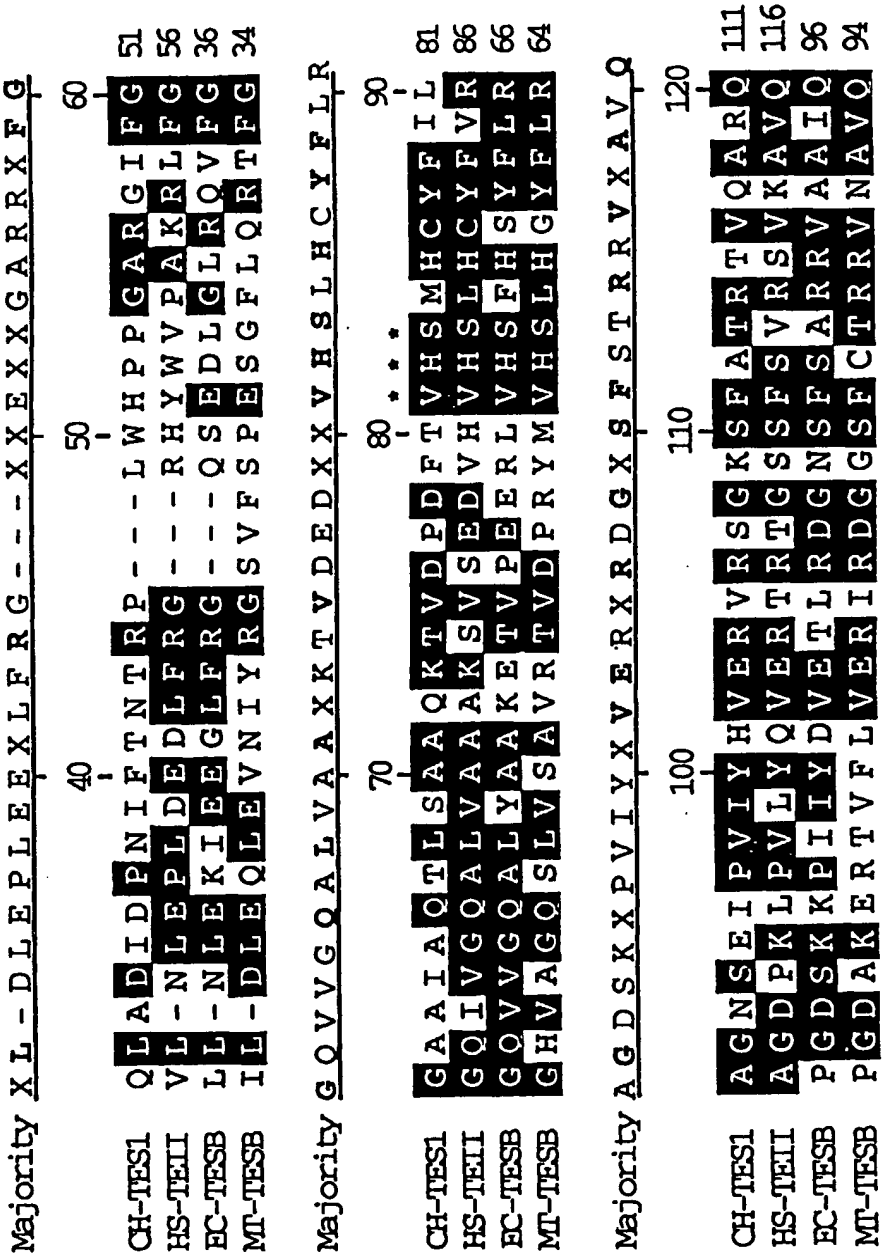

Majority XL-DLEPLEEXLFRG---XXEXXGARRXFG
CH-TES1 QLADIDPNIFTNTRP---LWHPPGARGIFG 51
HS-TEII VL-NLEPLDELFRG---RHYWVPAKRLFG 56
EC-TESB LL-NLEKIEEGLFRG---QSEDLGLRQVFG 36
MT-TESB IL-DLEQLEVNIYRG SVFSPESGFLQRTFG 34
Majority GQVVGQALVAAXKTVDEDXXVHSLHCYFLR
CH-TES1 GAAIAQTLSAAQKTVDPDFTVHSMHCYFIL 81
HS-TEII GQIVGQALVAAAKSVSEDVHVHSLHCYFVR 86
EC-TESB GQVVGQALYAAKETVPEERLVHSFHSYFLR 66
MT-TESB GHVAGQSLVSAVRTVDPRYMVHSLHGYFLR 64
Majority AGDSKXPVIYXVEERXRDGXSFSTRRVXAVQ
CH-TES1 AGNSEIPVIYHVERVRSGKSFATRTVQARQ 111
HS-TEII AGDPKLPVLYQVERTRTGSSFSVRSVKAVQ 116
EC-TESB PGDSKKPIIYDVETLRDGNSFSARRVAAIQ 96
MT-TESB PGDAKERTVFLVERIRDGGSFCTRRVNAVQ 94

FIG. 22A

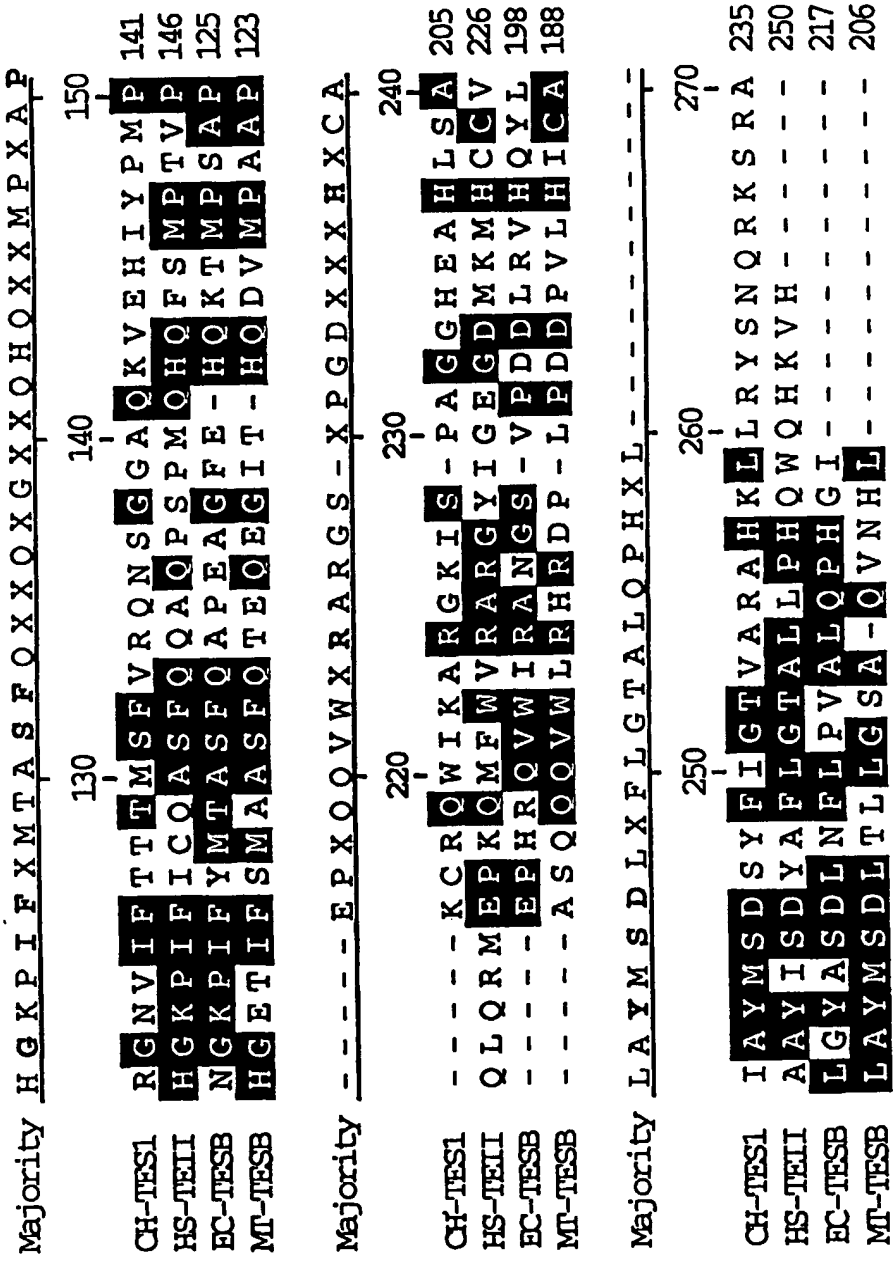

Majority HGKPIFXMTASFQXXQXGXXQHQXXMPXAP
130
140
150
CH-TES1 RGNVIFTTMSFVRQNSGGAQKVEHIYPMP 141
HS-TEII HGKPIFICQASFQQAQPSPMQHQFSMPTVP 146
EC-TESB NGKPIFYMTASFQAPEAGFE-HQKTMPSAP 125
MT-TESB HGETIFSMASFQTEQEGIT-HQDVMPAAP 123
Majority ----EPXQQVWXRARGS-XPGDXXXXHXCA
220
230
240
CH-TES1 -----KCRQWIKARGKIS-PAGGHEAHLSA 205
HS-TEII QLQRMEPKQMFWVRARGYIGEGDMKMHCCV 226
EC-TESB ----EPHRQVWIRANGS-VPDDLRVHQYL 198
MT-TESB ----ASQQVWLRHRDP-LPDDPVLHICA 188
Majority LAYMSDLXFLGTALQPHXL----------
250
260
270
CH-TES1 IAYMSDSYFIGTVARAHKLLRYSNQRKSRA 235
HS-TEII AAYISDYAFLGTALLPHQWQHKVH---- 250
EC-TESB LGYASDLNFLPVALQPHGI---------- 217
MT-TESB LAYMSDLTLLGSA-QVNHL---------- 206

FIG. 22B

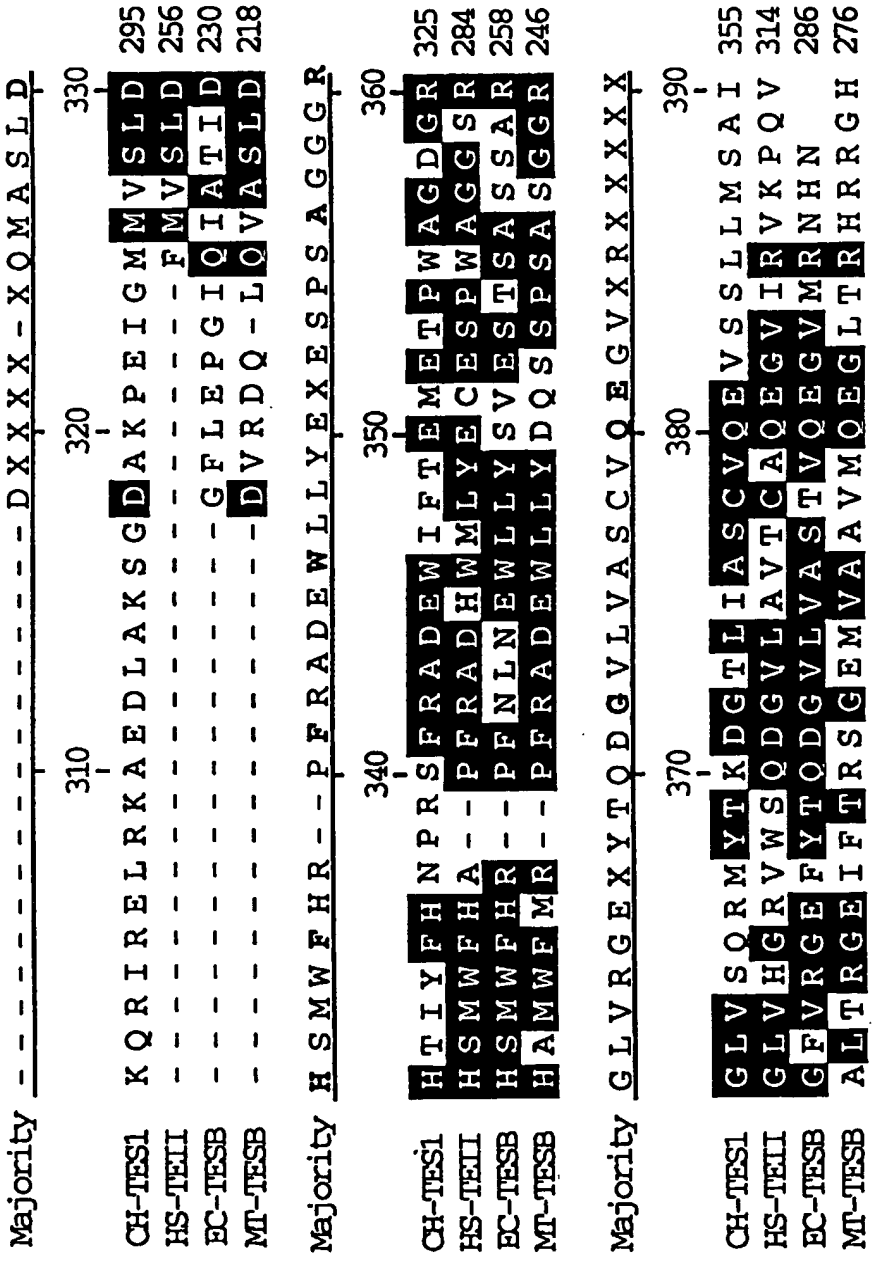

Majority
CH-TES1
HS-TEII
EC-TESB
MT-TESB
310
320
330
340
350
360
370
380
390
295
256
230
218
325
284
258
246
355
314
286
276

FIG. 22C

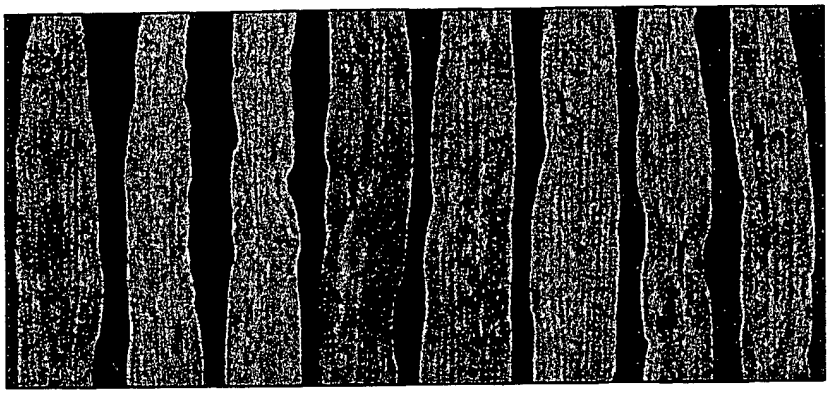


FIG. 23A

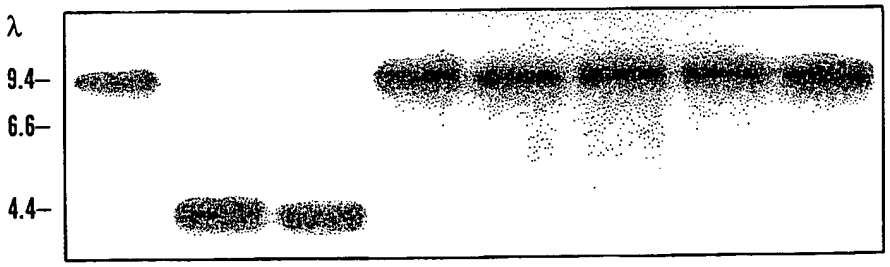


FIG. 23B

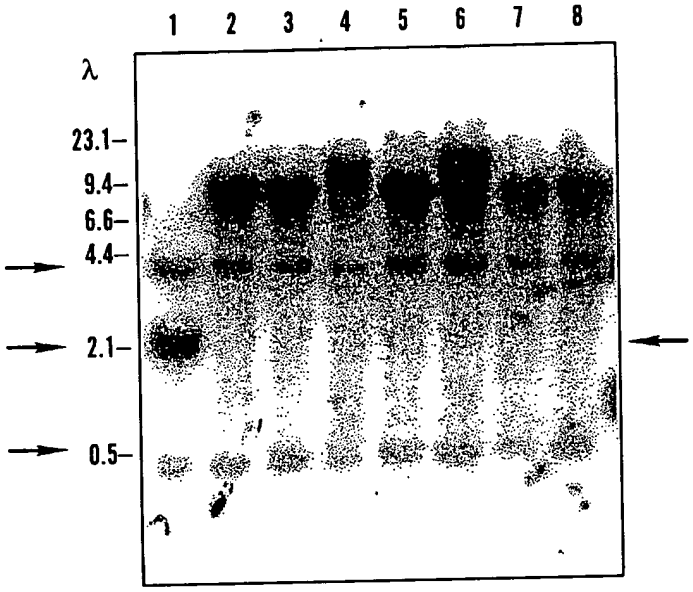


FIG. 24

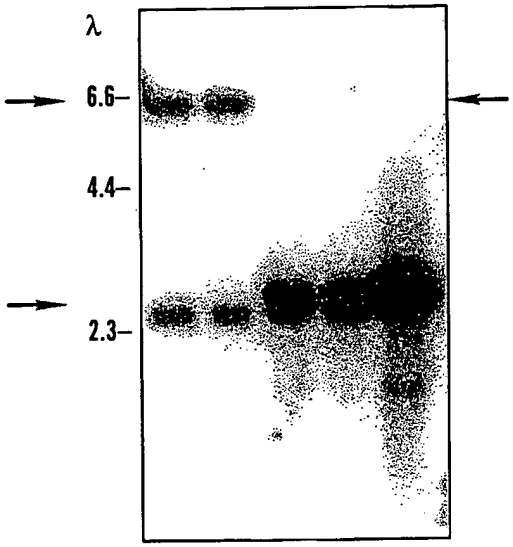


FIG. 25

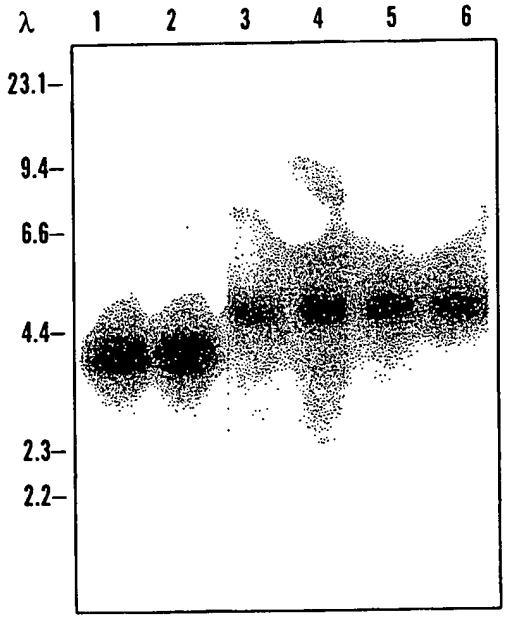


FIG. 26

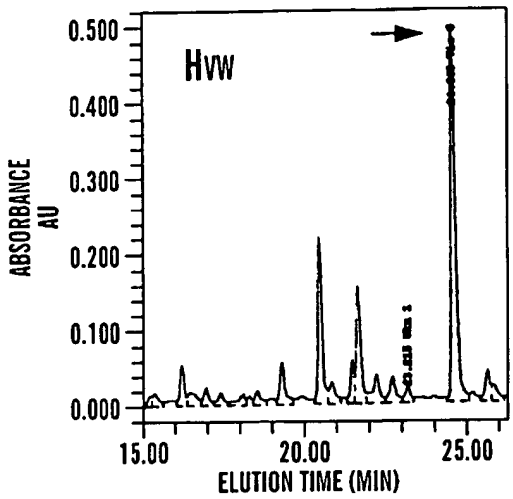


FIG. 27A

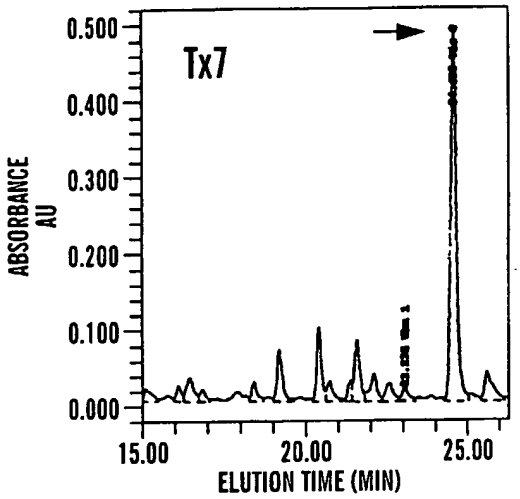


FIG. 27B

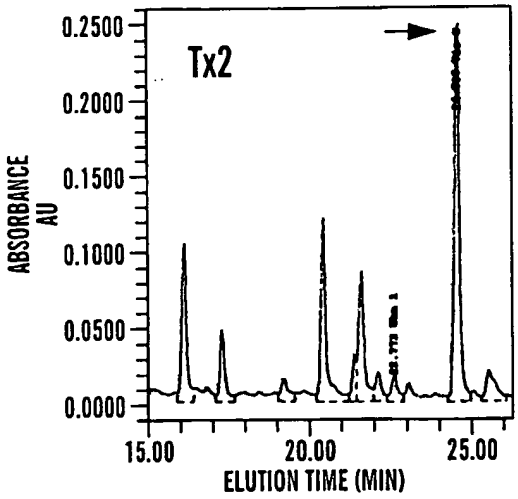


FIG. 27C

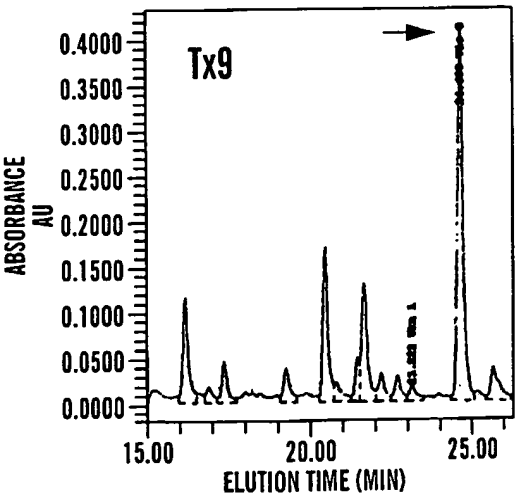


FIG. 27D

FIG. 28A

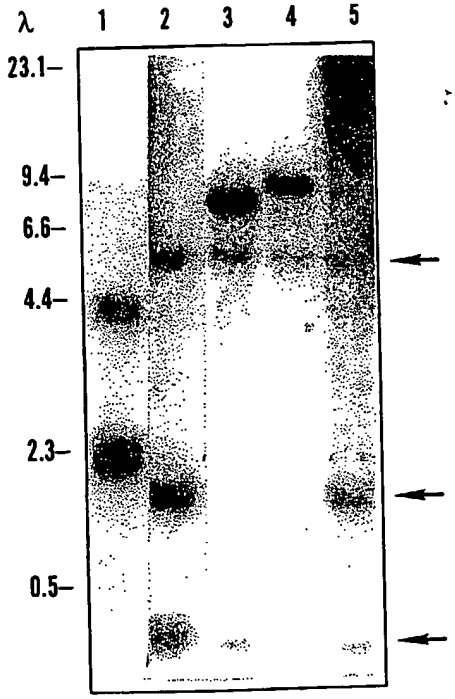


FIG. 28B

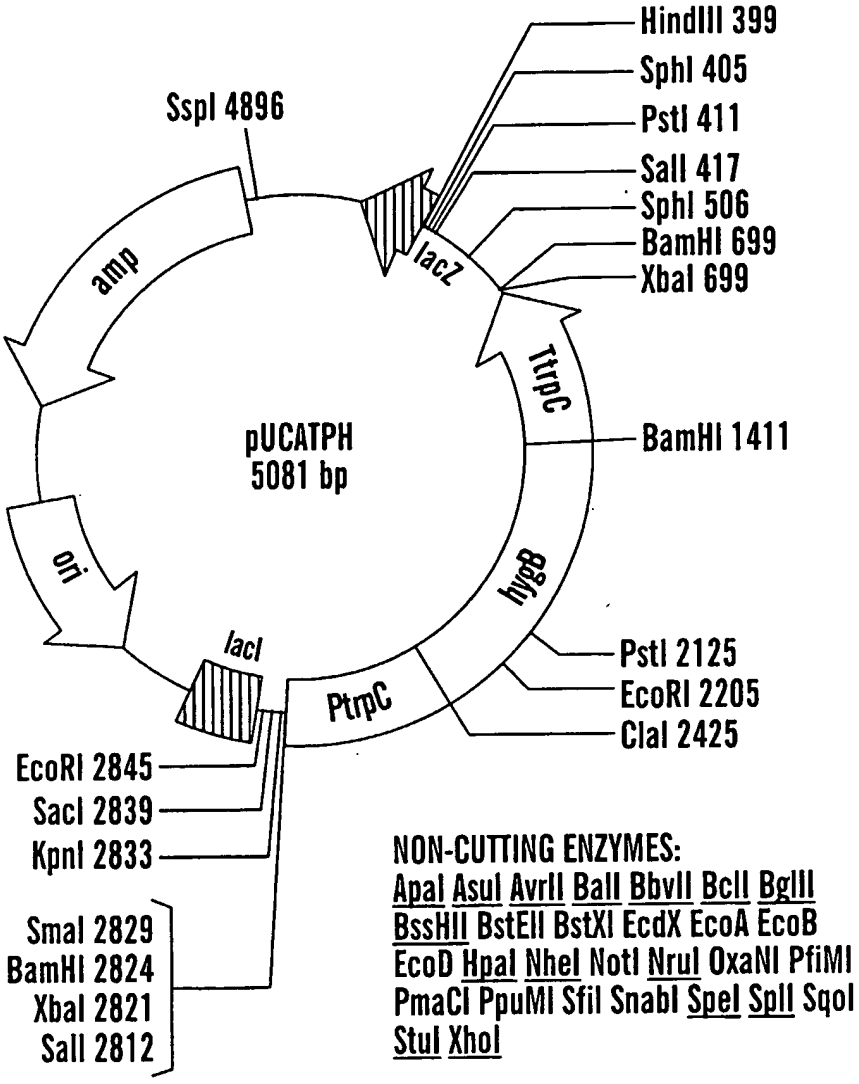

HindIII 399
SphI 405
PstI 411
SalI 417
SphI 506
BamHI 699
XbaI 699
SspI 4896
lacZ
amp
TtrpC
BamHI 1411
pUCATPH
5081 bp
ori
hygB
lacI
PtrpC
PstI 2125
EcoRI 2205
ClaI 2425
EcoRI 2845
SacI 2839
KpnI 2833
SmaI 2829
BamHI 2824
XbaI 2821
SalI 2812
NON-CUTTING ENZYMES:
ApaI AsuI AvrII BalI BbvII BclI BglII
BssHII BstEII BstXI EcdX EcoA EcoB
EcoD HpaI NheI NotI NruI OxaNI PfiMI
PmaCI PpuMI SfiI SnaBI SpeI SplI SqoI
StuI XhoI

FIG. 29

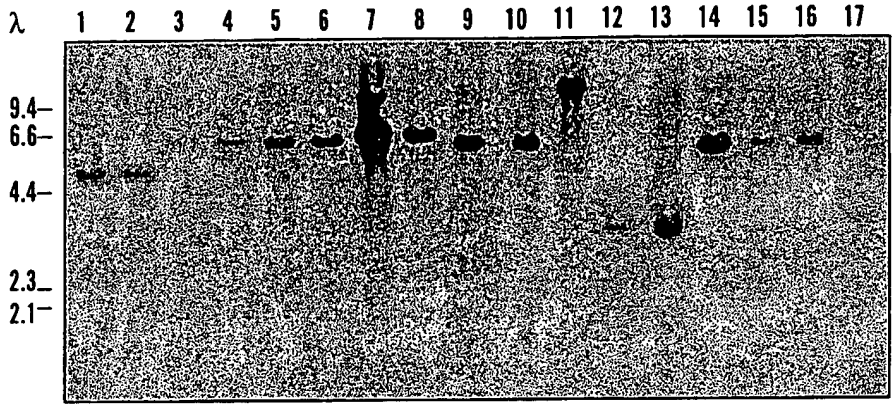


FIG. 30A

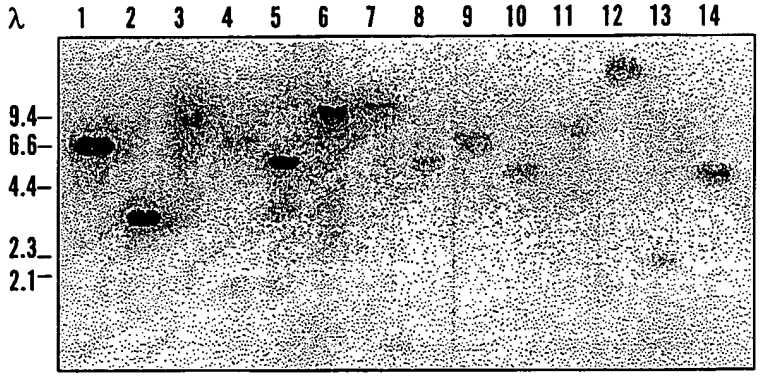


FIG. 30B

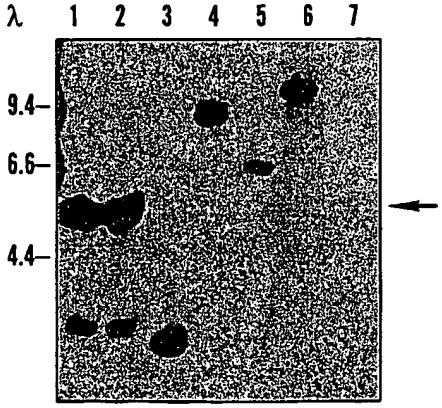


FIG. 30C

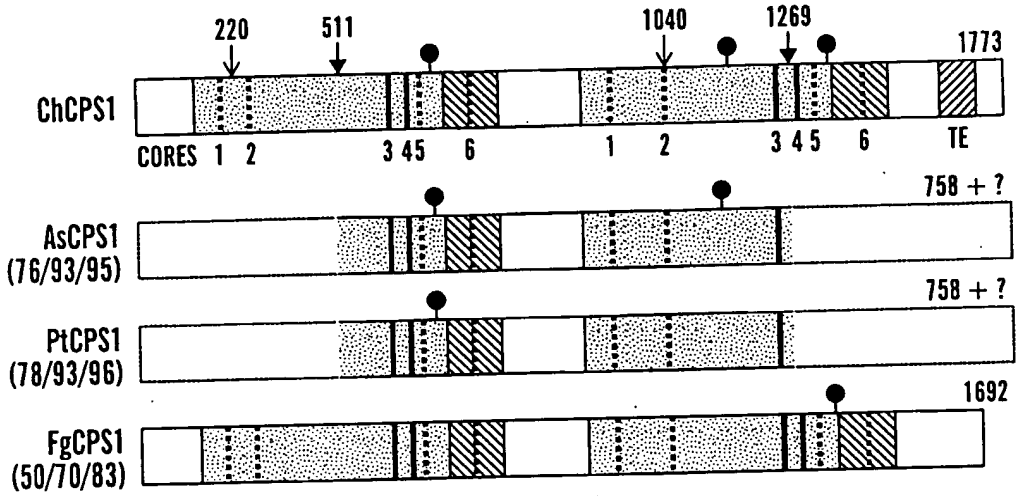

220
511
1040
1269
1773
ChCPS1
CORES 1 2
3 4 5
6
1
2
3 4 5
6
TE
758 + ?
AsCPS1
(76/93/95)
758 + ?
PtCPS1
(78/93/96)
1692
FgCPS1
(50/70/83)

FIG. 31A

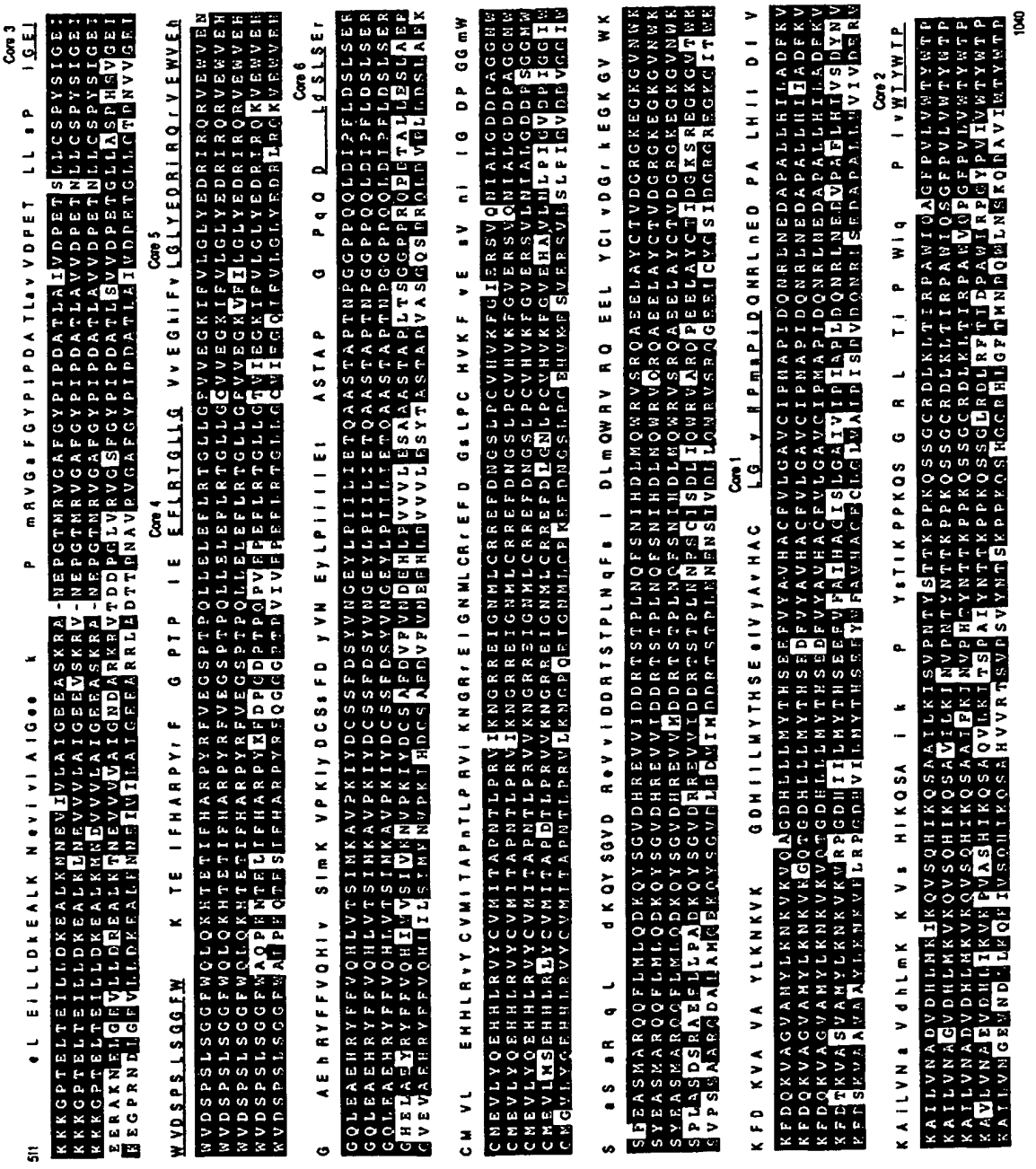


FIG. 31B

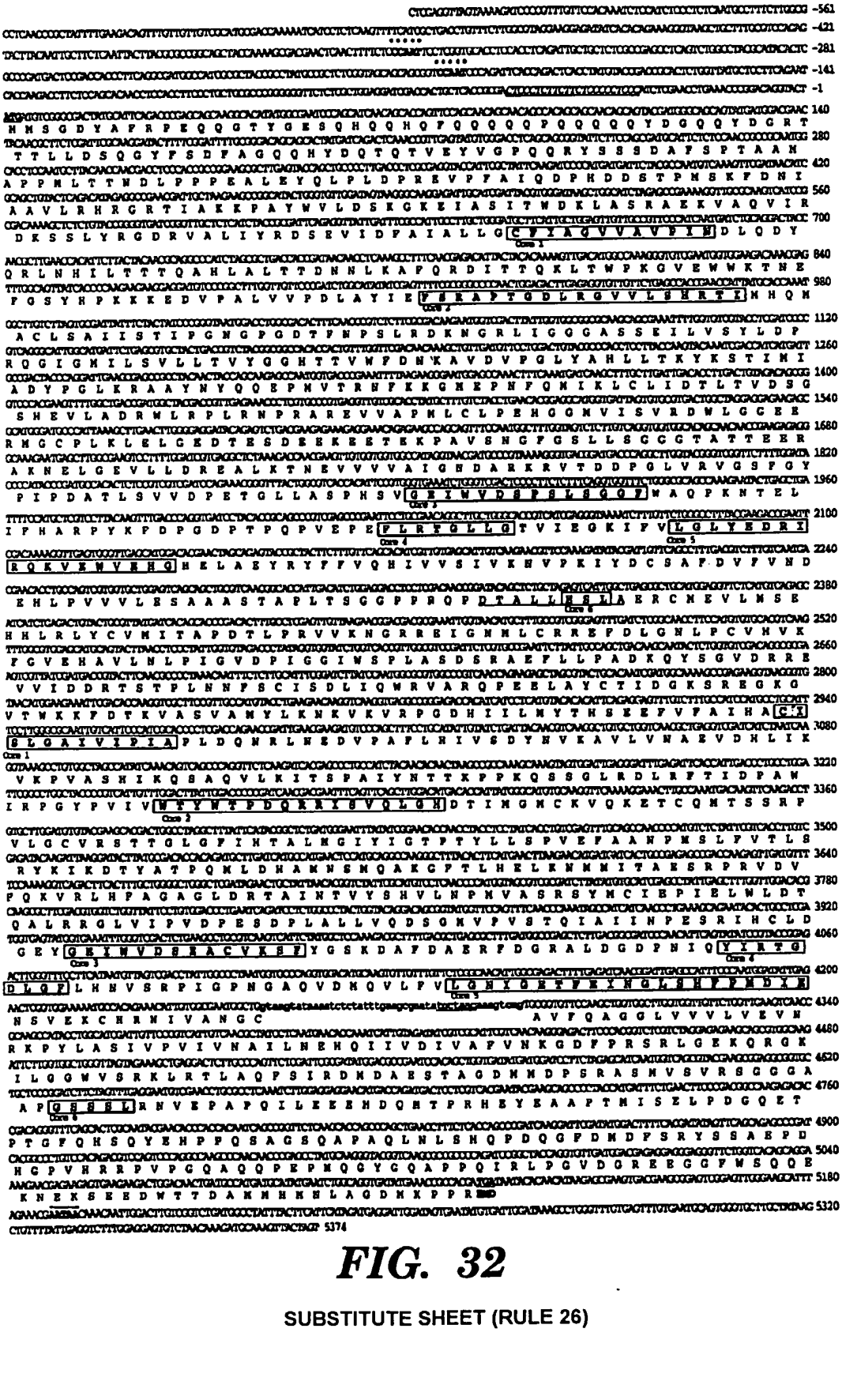


FIG. 32

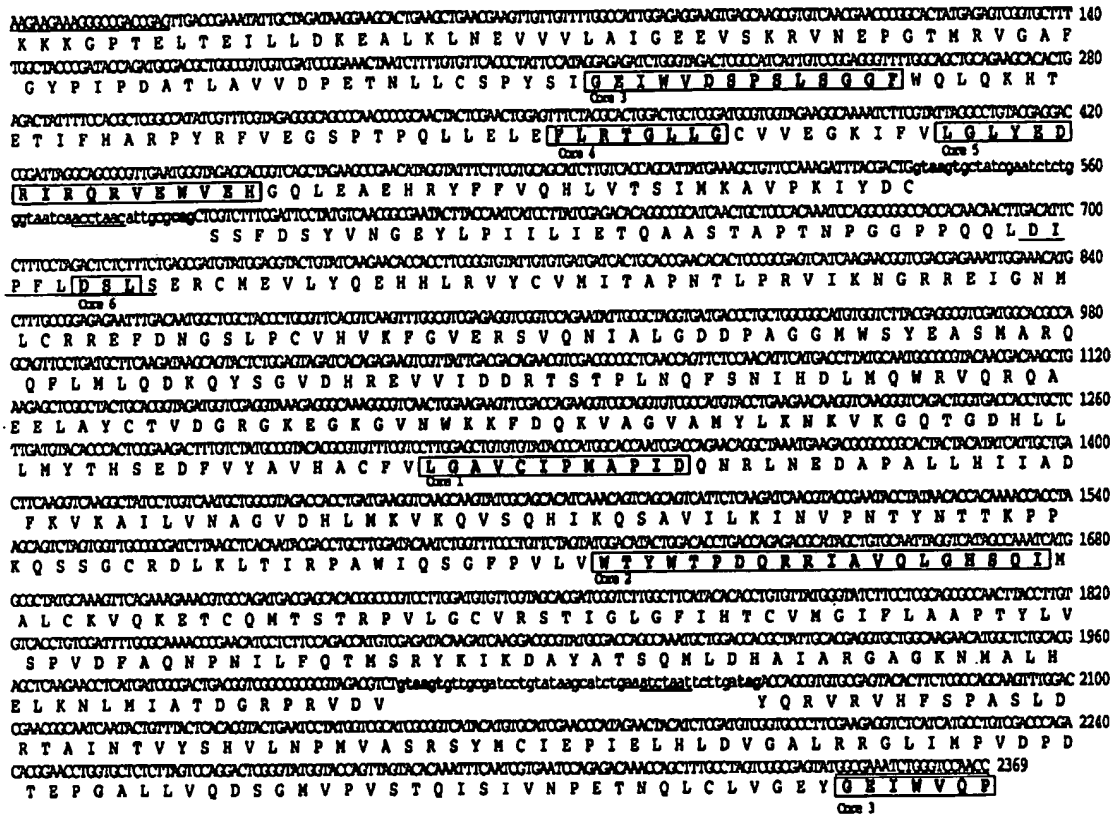


FIG. 33

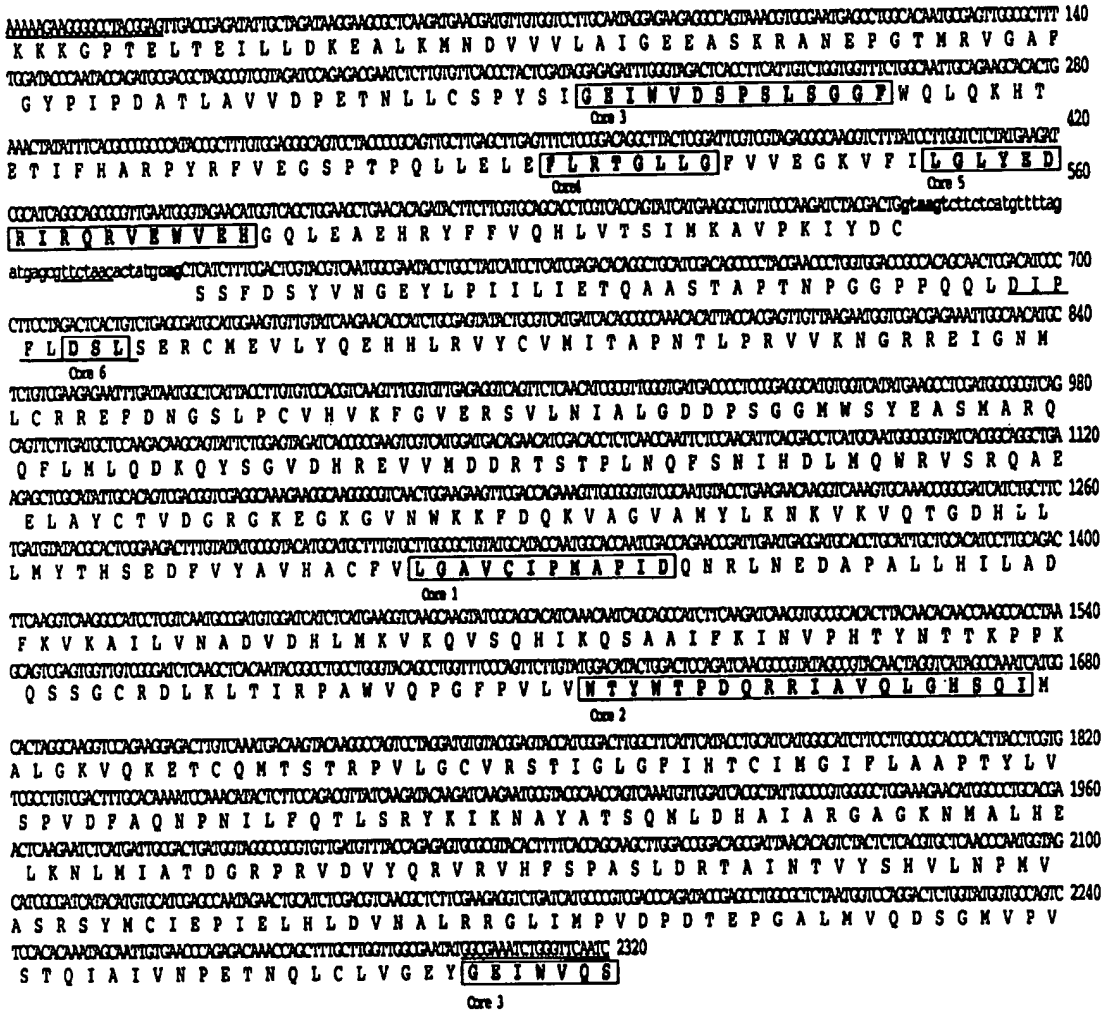


FIG. 34

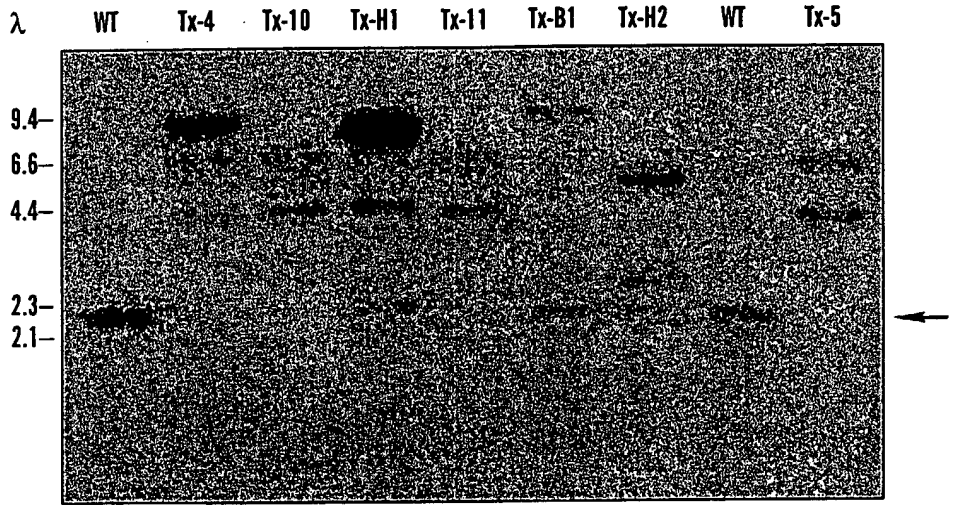


FIG. 35A

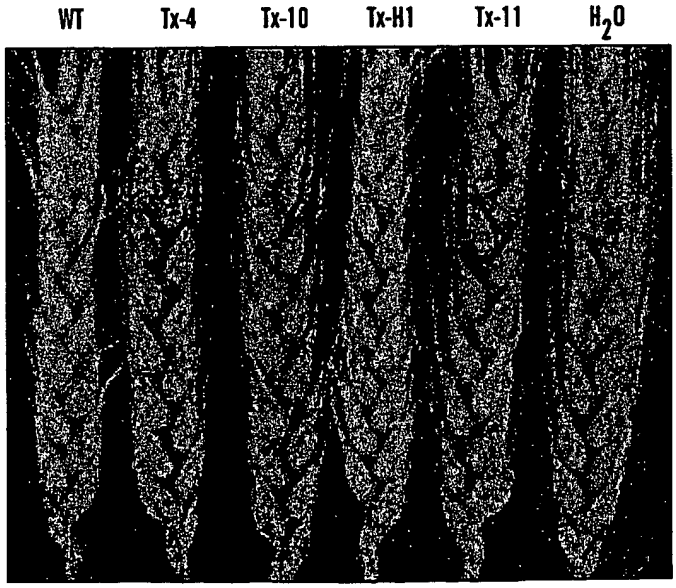


FIG. 35B

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 May 2001 (31.05.2001)

PCT

(10) International Publication Number
WO 01/38489 A3

(51) International Patent Classification[7]: A01H 5/00, C12N 5/14, 15/29, 15/52, 15/82

(21) International Application Number: PCT/US00/32227

(22) International Filing Date:
22 November 2000 (22.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/448,215 23 November 1999 (23.11.1999) US

(71) Applicant: **CORNELL RESEARCH FOUNDATION, INC.** [US/US]; Suite 105, 20 Thornwood Drive, Ithaca, NY 14850 (US).

(72) Inventors: **YODER, Olen, C.**; 44 Hungerford Road, Ithaca, NY 14850 (US). **TURGEON, Barbara, C.**; 608 Mitchell Street, Ithaca, NY 14850 (US). **LU, Shun-Wen**; 604 Winston Court, Apt. #4, Ithaca, NY 14850 (US).

(74) Agents: **GOLDMAN, Michael, L.** et al.; Nixon Peabody LLP, Clinton Square, P.O. Box 31051, Rochester, NY 14603 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
31 January 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/38489 A3

(54) Title: PEPTIDE SYNTHETASE GENE CPS1

(57) Abstract: The present invention relates to genes cloned from the plant pathogens *Cochliobolus heterostrophus, Alternaria solani, Fusarium graminearium,* and *Pyrenophora teres*, that encode a CPS1 peptide synthetase required for fungal pathogenesis. The nucleic acid molecules in a vector, a host cell, or a plant is also disclosed. The invention further provides a protein or polypeptide encoded by the CPS1 genes. Other aspects of the invention relate to a method of imparting disease resistance to a plant by overexpressing a protein of the present invention in a plant and a method for identifying inhibitors of a CPS1 protein in a sample.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/32227

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :A01H 5/00; C12N 5/14, 15/29, 15/52, 15/82

US CL :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1, 412, 419, 468; 536/23.2, 23.6; 800/279, 306, 314, 315, 316, 317.2, 317.4, 320, 320.1, 320.2, 320.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN, AGRICOLA, CAPLUS, BIOSIS, EMBASE, USPAT
search terms: peptide synthase, peptide synthetase, DNA, cDNA, gene, nucleic, disease, fungal, bacterial, pathogen, resistance

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NIKOLSKAYA et al. Identification of Peptide Synthetase-Encoding Genes from Filamentous Fungi Producing Host-Selective Phytotoxins or Analogs. Gene. 1995. Vol. 165, pages 207-211.	49-53

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 23 MAY 2001	Date of mailing of the international search report 26 JUL 2001
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer AMY NELSON Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/32227

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 1-5,25-44
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

applicant did not submit a computer readable form of the sequence listing, and hence SED ID NO:41 could not be searched.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-5,25-44,49-53

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/32227

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/320.1, 412, 419, 468; 536/23.2, 23.6; 800/279, 306, 314, 315, 316, 317.2, 317.4, 320, 320.1, 320.2, 320.3

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

Group I, claim(s)1-5,25-44,49-53, drawn to DNA of SEQ ID NO:41, vector, transformed host cell, transgenic plant, and plant transformation method. Claims 25-44 and 49-53 will be examined to the extent they read on SEQ ID NO:41.
Group II, claim(s) 6-8, drawn to protein of SEQ ID NO:42.
Group III, claim(s) 9-13, 25-44,49-53, drawn to DNA of SEQ ID NO:43, vector, transformed host cell, transgenic plant, and plant transformation method. Claims 25-44 and 49-53 will be examined to the extent they read on SEQ ID NO:43.
Group IV, claim(s) 14-16, drawn to protein of SEQ ID NO:44.
Group V, claim(s) 17-21,25-44,49-53, drawn to DNA of SEQ ID NO:45, vector, transformed host cell, transgenic plant, and plant transformation method. Claims 25-44 and 49-53 will be examined to the extent they read on SEQ ID NO:45.
Group VI, claim(s) 22-24, drawn to protein of SEQ ID NO:46.
Group VII, claim(s) 45-48, drawn to inhibitor selection method.

The inventions listed as Groups I, II, III, IV, V, VI and VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The phrase "CPS1" and "stringent conditions" are not well defined in the disclosure, and hence the claims of Group I encompass essentially any DNA from a plant pathogen encoding a cyclic peptide synthetase. Nikolskaya (Gene 165: 207-211, 1995) teaches various CPS genes isolated from plant pathogens (see entire article). Hence, there is no special technical feature under PCT Rule 13.2 which links the DNA of Group I with the protein of Group II.

The DNAs of Group III and V, and the proteins of Group IV and VI differ in composition and structure from the DNA of Group I and the protein of Group II, respectively, and hence are not so linked by a special technical feature. Separate searches and considerations would be required for examination of each of the nucleic acid sequences or amino acid sequences.

Also, the inhibitor selection method of Group VII is distinct from the plant transformation method of Group I in starting materials, method steps, and end products.

Therefore, the inventions of Groups I, II, III, IV, V, VI and VII do not relate to a single inventive concept under PCT Rule 13.1.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/32227

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/320.1, 412, 419, 468; 536/23.2, 23.6; 800/279, 306, 314, 315, 316, 317.2, 317.4, 320, 320.1, 320.2, 320.3

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

Group I, claim(s)1-5,25-44,49-53, drawn to DNA of SEQ ID NO:41, vector, transformed host cell, transgenic plant, and plant transformation method. Claims 25-44 and 49-53 will be examined to the extent they read on SEQ ID NO:41.
Group II, claim(s) 6-8, drawn to protein of SEQ ID NO:42.
Group III, claim(s) 9-13, 25-44,49-53, drawn to DNA of SEQ ID NO:43, vector, transformed host cell, transgenic plant, and plant transformation method. Claims 25-44 and 49-53 will be examined to the extent they read on SEQ ID NO:43.
Group IV, claim(s) 14-16, drawn to protein of SEQ ID NO:44.
Group V, claim(s) 17-21,25-44,49-53, drawn to DNA of SEQ ID NO:45, vector, transformed host cell, transgenic plant, and plant transformation method. Claims 25-44 and 49-53 will be examined to the extent they read on SEQ ID NO:45.
Group VI, claim(s) 22-24, drawn to protein of SEQ ID NO:46.
Group VII, claim(s) 45-48, drawn to inhibitor selection method.

The inventions listed as Groups I, II, III, IV, V, VI and VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The phrase "CPS1" and "stringent conditions" are not well defined in the disclosure, and hence the claims of Group I encompass essentially any DNA from a plant pathogen encoding a cyclic peptide synthetase. Nikolskaya (Gene 165: 207-211, 1995) teaches various CPS genes isolated from plant pathogens (see entire article). Hence, there is no special technical feature under PCT Rule 13.2 which links the DNA of Group I with the protein of Group II.

The DNAs of Group III and V, and the proteins of Group IV and VI differ in composition and structure from the DNA of Group I and the protein of Group II, respectively, and hence are not so linked by a special technical feature. Separate searches and considerations would be required for examination of each of the nucleic acid sequences or amino acid sequences.

Also, the inhibitor selection method of Group VII is distinct from the plant transformation method of Group I in starting materials, method steps, and end products.

Therefore, the inventions of Groups I, II, III, IV, V, VI and VII do not relate to a single inventive concept under PCT Rule 13.1.